

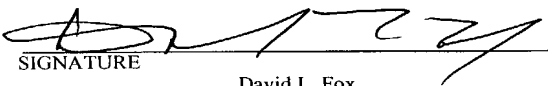
10018730 10018730

3009 Rec'd PCT/PTO 18 DEC 2001

FORM PTO 1390 (REV 9-2001)		U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER HO-P02353US0	
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371				U.S. APPLICATION NO. (If known, see 37 CFR 1.5) <div style="font-size: 1.5em; font-weight: bold;">10/018730</div>	
INTERNATIONAL APPLICATION NO. WO 00/78973 - 6800 2379		INTERNATIONAL FILING DATE 19/06/2000		PRIORITY DATE CLAIMED 18/06/1999	
TITLE OF INVENTION PROCESS FOR OXIDISING AROMATIC COMPOUNDS					
APPLICANT(S) FOR DO/EO/US Luet L. Wong, et al.					
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:					
<ol style="list-style-type: none"> 1. <input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371 2. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing 35 U.S.C. 371 3. <input type="checkbox"/> This is an express request to begin national examination procedures (35 U.S.C. 371 (f)). The submission must include items (5), (6), (9) and (21) indicated below. 4. <input type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (PCT Article 31). 5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371 (c)(2)) <ol style="list-style-type: none"> a. <input checked="" type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau). b. <input type="checkbox"/> has been communicated by the International Bureau. c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US). 6. <input type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C. 371 (c)(2)) <ol style="list-style-type: none"> a. <input type="checkbox"/> is attached hereto. b. <input type="checkbox"/> has been previously submitted under 35 U.S.C. 154(d)(4). 7. <input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3)) <ol style="list-style-type: none"> a. <input checked="" type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau). b. <input type="checkbox"/> have been communicated by the International Bureau. c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired. d. <input type="checkbox"/> have not been made and will not be made. 8. <input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c)(3)). 9. <input type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)). 10. <input type="checkbox"/> An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)). 					
Items 11 to 20 below concern document(s) or information included:					
<ol style="list-style-type: none"> 11. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98. 12. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included. 13. <input checked="" type="checkbox"/> A FIRST preliminary amendment. 14. <input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment. 15. <input type="checkbox"/> A substitute specification. 16. <input type="checkbox"/> A change of power of attorney and/or address letter. 17. <input checked="" type="checkbox"/> A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C. 1.821 - 1.825. 18. <input type="checkbox"/> A second copy of the published international application under 35 U.S.C. 154(d)(4). 19. <input type="checkbox"/> A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4). 20. <input checked="" type="checkbox"/> Other items or information: Power of Attorney 					
					EK102703248US

531 Rec'd PCT/

18 DEC 2001

U.S. APPLICATION NO. 107018730 UNKNOWN		INTERNATIONAL APPLICATION NO. 6800 WO 00/78973		ATTORNEY'S BOOKLET NUMBER 0379 HO-P02353US0	
21. <input checked="" type="checkbox"/> The following fees are submitted:				CALCULATIONS PTO USE ONLY	
BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)):					
<input type="checkbox"/> Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO \$1040.00					
<input checked="" type="checkbox"/> International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO . . . \$890.00					
<input type="checkbox"/> International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO . . . \$740.00					
<input type="checkbox"/> International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) \$710.00					
<input type="checkbox"/> International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) \$100.00					
ENTER APPROPRIATE BASIC FEE AMOUNT =				\$ 890.00	
Surcharge of \$ _____ for furnishing the oath or declaration later than				\$	
<input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492 (e)).					
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	18-20 =		x	\$	0.00
Independent claims	5-5 =		x	\$	0.00
MULTIPLE DEPENDENT CLAIM(s) (if applicable)			+	\$	
TOTAL OF ABOVE CALCULATIONS =				\$	890.00
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.				\$	
SUBTOTAL =				\$	890.00
Processing fee of \$ _____ for furnishing the English translation later than				\$	
<input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492 (f)). +					
TOTAL NATIONAL FEE =				\$	890.00
Fee for recording the enclosed assignment (37 CFR 1.21 (h)). Assignment must be accompanied by appropriate cover sheet (37 CFR 3.28, 3.31) (_____ per property).				+	\$
TOTAL FEES ENCLOSED =				\$	890.00
				Amount to be Refunded:	\$
				Charged:	\$
a. <input type="checkbox"/> A check in the amount of \$ _____ to cover the above fees is enclosed.					
b. <input checked="" type="checkbox"/> Please charge my Deposit Account No. <u>06-2375</u> in the amount of \$ <u>890.00</u> to cover the above fees. A duplicate copy of this sheet is enclosed.					
c. <input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge any additional fees which may be required or credit any overpayment to my Deposit Account No. <u>06-2375</u> . A duplicate copy of this sheet is enclosed.					
NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137 (a) or (b)) must be filed and granted to restore the application to pending status.					
SEND ALL CORRESPONDENCE TO:					
David L. Fox FULBRIGHT & JAWORSKI L.L.P. 1301 McKinney, Suite 5100 Houston, Texas 77010-3095 (713) 651-8231					
 SIGNATURE					
NAME					
40,612					
REGISTRATION NUMBER					

10/018730

PATENT

P02353US0

18 DEC 2001

531 Rec'd PC

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	§	
Luet Lok Wong	§	
Jonathan Peter Jones	§	
	§	Attorney Dkt: HO-P02353US0
Serial No.: Unassigned	§	
	§	Group Art Unit: Unassigned
Filed: Herewith	§	
	§	Examiner: Unassigned
For: PROCESS FOR OXIDISING	§	
AROMATIC COMPOUNDS	§	

PRELIMINARY AMENDMENT

BOX: PCT
 Commissioner for Patents
 Washington, DC 20231

Dear Sir:

This is a Preliminary Amendment to accompany the U.S. National Phase filing under 35 U.S.C. § 371 of PCT/GB00/02379, international filing date June 19, 2000, with priority claim to June 18, 1999 (GB 9914373.7). This Amendment substitutes the claims to be examined.

AMENDMENT

IN THE CLAIMS:

Please cancel claims 1-19 of the PCT application and substitute in their place the below-presented claims 1-18:

1. Process for oxidising a halo aromatic substrate which has more than one halogen atom, which process comprises oxidising said substrate with a monooxygenase enzyme, wherein a ring carbon of the substrate is oxidised.
2. Process according to claim 1 in which the enzyme comprises a substitution of an amino acid in the active site by an amino acid with a less polar side-chain.
3. Process according to claim 2 in which the enzyme comprises one or more other amino acid substitutions in the active site.

4. Process according to claim 1 in which the enzyme is:
 - (i) P450_{cam}, or
 - (ii) a naturally occurring homologue of (i), or
 - (iii) a mutant of (i) or (ii).
5. Process according to claim 4 in which the enzyme is one in which amino acid 96 of P450_{cam}, or the equivalent amino acid in a homologue, has been changed to an amino acid with a less polar side-chain.
6. Process according to claim 1 in which the halogen is chlorine.
7. Process according to claim 1 in which the aromatic compound is a benzene or biphenyl.
8. Process for oxidising a halo aromatic substrate, which process comprises oxidising said substrate with a monooxygenase enzyme, wherein the substrate is 1, 2-dichlorobenzene, 1, 2, 4-trichlorobenzene, 3,3'-dichlorobiphenyl, 2,2',4,5,5'-pentachlorobiphenyl, pentachlorobenzene or hexachlorobenzene
9. Process according to claim 8 in which the enzyme is:
 - (i) P450_{cam}, or
 - (ii) a naturally occurring homologue of (i), or
 - (iii) a mutant of (i) or (ii).
10. Process according to claim 1 which is carried out in a cell that expresses:
 - (a) a monooxygenase enzyme;
 - (b) an electron transfer reductase; and
 - (c) an electron transfer redoxin.
11. Process according to claim 10 in which (b) is putidaretoxin reductase or a homologue; or a fragment thereof; and/or (c) is putidaretoxin or a homologue; or a fragment thereof.
12. Process according to claim 10 wherein the cell is one in which the enzyme (a) does not naturally occur.
13. Process for oxidising a halo aromatic substrate which has more than one halogen atom, which process comprises oxidising said substrate with a monooxygenase enzyme, wherein a ring carbon of the substrate is oxidised, the process being carried out in a cell that expresses:
 - (a) a monooxygenase enzyme;
 - (b) an electron transfer reductase; and
 - (c) an electron transfer redoxin,

wherein the cell is one which in its naturally occurring form is able to oxidise a halo aromatic substrate.

14. A cell as defined in claim 13.

15. A transgenic non-human animal or transgenic plant whose cells are as defined in claim 10.
16. Method of treating a locus contaminated with a halo aromatic substrate comprising contacting the locus with:
- (i) a monooxygenase enzyme, or
 - (ii) a cell that expresses:
 - (a) a monooxygenase enzyme;
 - (b) an electron transfer reductase; and
 - (c) an electron transfer redoxin, or
 - (iii) a non-human transgenic animal or transgenic plant whose cells express (a), (b) and (c).
17. Process for selecting a mutant of a monooxygenase enzyme for its ability to oxidise a halo aromatic substrate, which process comprises screening a library of monooxygenase mutants for their oxidation effect on the substrate.
18. Process for oxidising a halo aromatic substrate which has more than one halogen atom, which process comprises oxidising said substrate with a monooxygenase enzyme, wherein a ring carbon of the substrate is oxidised, and wherein the enzyme is one that has been selected in a process according to claim 17.

REMARKS

Please substitute and examine the above-presented claims for those pending in the PCT application. Please begin numbering with the number 1. The above-presented claims represent re-written claims of the PCT application in U.S. format. No new matter has been added.

Sequence Listing

A Sequence Listing is submitted herewith in computer readable form and as a print out. The sequences are identical to those in the PCT application, and the content of the computer readable disk is identical to that of the prior out.

Please charge any fees due or credit any overpayment to the standing account of Fulbright & Jaworski L.L.P. No. 06-2375 under order no. P02353US0 on which the undersigned is authorized to sign..

CLAIMS

1. Process for oxidising a halo aromatic substrate which has more than one halogen atom, which process comprises oxidising said substrate with a monooxygenase enzyme, wherein a ring carbon of the substrate is oxidised.
2. Process according to claim 1 in which the enzyme comprises a substitution of an amino acid in the active site by an amino acid with a less polar side-chain.
3. Process according to claim 2 in which the enzyme comprises one or more other amino acid substitutions in the active site.
4. Process according to claim 1 in which the enzyme is:
 - (i) P450_{cam}, or
 - (ii) a naturally occurring homologue of (i), or
 - (iii) a mutant of (i) or (ii).
5. Process according to claim 4 in which the enzyme is one in which amino acid 96 of P450_{cam}, or the equivalent amino acid in a homologue, has been changed to an amino acid with a less polar side-chain.
6. Process according to claim 1 in which the halogen is chlorine.
7. Process according to claim 1 in which the aromatic compound is a benzene or biphenyl.
8. Process for oxidising a halo aromatic substrate, which process comprises oxidising said substrate with a monooxygenase enzyme, wherein the substrate is 1, 2-dichlorobenzene, 1, 2, 4-trichlorobenzene, 3,3'-dichlorobiphenyl, 2,2',4,5,5'-pentachlorobiphenyl, pentachlorobenzene or hexachlorobenzene
9. Process according to claim 8 in which the enzyme is:
 - (i) P450_{cam}, or
 - (ii) a naturally occurring homologue of (i), or
 - (iii) a mutant of (i) or (ii).
10. Process according to claim 1 which is carried out in a cell that expresses:
 - (a) a monooxygenase enzyme;
 - (b) an electron transfer reductase; and
 - (c) an electron transfer redoxin.

18. Process for oxidising a halo aromatic substrate which has more than one halogen atom, which process comprises oxidising said substrate with a monooxygenase enzyme, wherein a ring carbon of the substrate is oxidised, and wherein the enzyme is one that has been selected in a process according to claim 17.

I hereby certify that this correspondence is being deposited with the U S Postal Service as Express Mail, Airbill No.EU098493435US, in an envelope addressed to Box PCT, Commissioner for Patents, Washington, DC 20231, on the date shown below

Dated April 4, 2002

Signature:

Linda Bourg
(Linda Bourg)

13 Rec'd PCT/PTO 04 APR 2002
10/018730

Docket No.: P02353US1
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
Wong et al.

Application No.: 10/018,730
Filed: December 18, 2001

Group Art Unit: Not yet assigned

Int'l Appln. No.: PCT/GB00/02379
Filed: June 19, 2000

Examiner: Not yet assigned

For: PROCESS FOR OXIDISING AROMATIC
COMPOUNDS

SECOND PRELIMINARY AMENDMENT

Box PCT, DOEO
Commissioner for Patents
Washington, DC 20231

Dear Sir:

Prior to examination on the merits, please amend the above-identified U.S. patent application as follows:

1. **Attorney Docket Number:**

Please note that the attorney docket number has been changed from P02353US0 to: **P02353US1**. Applicants respectfully request that this change be reflected on all correspondence from the Office.

2. **In the Specification:**

Please substitute the below amended paragraphs to correct typographical errors, to direct SEQ ID NOs or corrections thereof into the specification.

Paragraph on page 6, at lines 2-7:

Typically (i) is any allelic variant of P450_{cam} of *Pseudomonas putida* (e.g. of the polypeptide sequence shown in SEQ ID No. 2). Typically (ii) is a species homologue of (i) which has sequence homology with (i), and is typically P450_{BM-3} of *Bacillus megaterium* (e.g. the polypeptide sequence shown in SEQ ID No. 4 and the nucleotide sequence is shown in SEQ ID NO: 3), P450_{terp} of *Pseudomonas sp.*, P450_{eryF} of *Saccharopolyspora erythraea*, or P450 105 D1 (CYP105) of *Streptomyces griseus* strains.

Paragraph on page 6, at lines 22-32:

The discussion below provides examples of the positions at which substitutions may be made in P450_{cam}. The same substitutions may be made at equivalent positions in the homologues. Standard nomenclature is used to denote the mutations. The letter of the amino acid present in the natural form is followed by the position, followed by the amino acid in the mutant. To denote multiple mutations in the same protein each mutation is listed separated by hyphens. The mutations discussed below using this nomenclature specify the natural amino acid in P450_{cam}, but it is to be understood that the mutation could be made to a homologue which has a different amino acid at the equivalent position. Note that the amino acid numbering shown in SEQ ID No. 2 for P450_{cam} does not correspond to the numbering used in the description to denote mutations. The numbering in SEQ ID No. 2 is one more than the numbering in the description for a particular position.

Paragraph on page 9, at lines 4-10:

Alternatively a linker may be present between the components. The linker generally comprises amino acids that do not have bulky side chains and therefore do not obstruct the folding of the protein subunits. Preferably the amino acids in the linker are uncharged. Preferred amino acids in the linker are glycine, serine, alanine or threonine. In one embodiment the linker comprises the sequence N-Thr-Asp-Gly-Gly-Ser-Ser-Ser-C (SEQ ID NO:6). The linker is typically from at least 5 amino acids long, such as at least 10, 30 or 50 or more amino acids long.

Paragraph on page 11, at lines 2-6:

In one embodiment the three genes encoding the three proteins of the P450_{cam} system, i.e. camA, camB, and camC are placed in the mobile regions of standard transposon vectors and incorporated into the genome of *Pseudomonas* and flavobacteria. Alternatively plasmid vectors for expressing these genes may be used, in which case the P450_{cam} gene cluster will be extra-chromosomal.

Oligonucleotide-directed site-specific mutagenesis was carried out by the Kunkel method (Kunkel, T. A. *Proc. Natl. Acad. Sci. USA* 1985, **82**, 488-492) using the Bio-Rad Mutagen kit. The recommended procedure is summarised as follows. An M13 mp19 subclone of the *camC* gene encoding P450_{cam} (SEQ ID NO: 1) was propagated in the *E. coli* strain CJ236. This strain has the *dut*⁻*ung*⁻ phenotype and thus will tolerate the inclusion of uracil in place of thymine in DNA molecules. After three cycles of infection, uracil-containing single stranded (USS) M13 DNA was readily isolated by phenol extraction of mature M13 phage particles excreted into the growth medium. The mutagenic oligonucleotide (or oligonucleotides) were phosphorylated with T4 polynucleotide kinase and then annealed to the USS template. The four nucleotides, DNA polymerase, DNA ligase, ATP and other chemical components were added and the second strand was synthesised *in vitro*. The double stranded form thus obtained was transformed into the *dut*⁺*ung*⁺ *E. coli* strain MV1190, which should degrade the uracil-containing template strand and propagate the mutant strand synthesised *in vitro*. Plaques were picked and phages of possible mutants grown in *E. coli* strains MV1190 or TG1. The single-stranded DNA from these were sequenced to determine whether the mutagenesis reaction was successful. The mutagenic efficiency was 50 - 80%.

For the *camA* gene the primer below (SEQ ID NO: 5) was used at the 5' end of the gene to introduce the *Eco* RI cloning site and to change the first codon from GTG to the strong start codon ATG.

5'- GAG ATT AAG AAT TCA TAA ACA CAT GGG AGT GCG TGC CAT ATG AAC
GCA AAC

Eco RI RBS | → *camA*

At the 3' end of *camA* the primer (nucleotide sequence is SEQ ID NO: 7) was designed such that 15 bases are complementary to nucleotide sequence of the last five amino acid residues of *camA*. The stop codon immediately after the GCC codon for the last amino acid was removed, and then part of a seven amino acid linker (Thr Asp Gly Gly Ser Ser Ser; SEQ ID NO: 6) which contained a *Bam* HI cloning site (GGATCC = Gly Ser) was introduced. The coding sequence was thus (amino acid sequence is SEQ ID NO:8):

5'- GAA CTG AGT AGT GCC ACT GAC GGA **GGA TCC** TCA TCG-3'
camA → Thr Asp Gly Gly Ser
 | *Bam* HI |

Paragraph on page 17, at lines 5-6:

The primer sequence shown below (SEQ ID NO: 9) is the reverse complement used for PCR:

5'- CGA TGA **GGA TCC** TCC GTC AGT GGC ACT ACT CAG TTC-3'

Paragraph on page 17, at lines 8-13:

For the *camB* gene the primer at the 5' end (nucleotide sequence is SEQ ID NO: 10; amino acid sequence is SEQ ID NO: 11) incorporated the second half of the peptide linker between the reductase and redoxin proteins, and the restriction site *Bam* HI for joining the two amplified genes together.

5'- TCA TCG **GGA TCC** TCA TCG **ATG TCT** AAA GTA GTG TAT-3'

Gly Ser Ser Ser |→ *camB*
|*Bam* HI| Start

Paragraph on page 17, at lines 14-19:

At the 3' end of *camB* the primer incorporates 12 nucleotides complementary to the end of *camB* followed by the stop codon TAA, a 6 nucleotide spacer before the GGAG ribosome binding site. *Xba* I and *Hind* III sites were then added to allow cloning of the *camC* gene when required. The sequence of the coding strand was therefore (SEQ ID NO: 12):

5'- CCC GAT AGG CAA TGG **TAA** TCA TCG **GGAG TCT** AGA GCA TCG **AAG CTT**
TCA TCG-3'

CamB →|stop RBS *Xba* I *Hind* III

Paragraph on page 17, at lines 20-21:

The primer shown below (SEQ ID NO: 13) is the reverse complement used for PCR:

5'-CGA TGA AAG CTT CGA TGC TCT AGA CTCC CGA TGA TTA CCA
TTG CCT ATC GGG -3'

Paragraph on page 19, at lines 2-6:

We used the internal and unique restriction site *Mlu* I (recognition sequence ACGCGT) within the *camB* gene as the starting point so that the PCR product has a different size from the PCR template fragment. The primers were as follows:

5'- TCA TCG **ACG CGT** CGC GAA CTG CTG-3' (SEQ ID NO: 14)

where the *Mlu* I site is in bold.

Application No. 10/018,730

Docket No. HO-P02353US1

Paragraph on page 19, at lines 7-10:

The desired coding sequence at the 3' end of the *camB* gene (SEQ ID NO:15) was:

5'- CCC GAT AGG CAA TGG **TAA** GTA GGT GAA TAT CTA ATC CCC ATC
camB →|stop

TAT GCG CGA GTG **GAG TCT** AGA GTT CGA-3'
RBS *Xba* I

Paragraph bridging page 20, lines 25 to page 21, line 7:

The primers shown below maintain the *Hind* III cloning site AAGCTT:

5'-TCA TCG **AAG CTT** GGC TGT TTT-3' (SEQ ID NO: 16)
Hind III |→ vector

At the other end the coding sequence desired was (SEQ ID NO: 17):

5'-ACA ATT TCA CAC **AGGA TCT** AGA C **CAT ATG** TCA TCG **AAG CTT**
TCA TCG-3'
Vector →|RBS *Xba* I *Nde* I *Hind* III

This sequence maintained the *Nde* I and *Hind* III sites but the new *Xba* I site was introduced upstream of the *Nde* I site. The PCR primer used was the reverse complement of the desired sequence (SEQ ID NO: 18):

5'-CGA TGA AAG CTT CGA TGA CAT ATG GTC T AGA TCCT GTG
TGA AAT TGT-3'

Application No. 10/018,730

Docket No. HO-P02353US1

REMARKS

The Action, mailed February 4, 2002, requires that a substitute Sequence Listing that is in compliance with 37 C.F.R. §§ 1.822 and/or 1.823 be provided. In response, Applicants have provided herewith a substitute copy of the corrected Sequence Listing in computer readable form and in paper form under 37 C.F.R. § 1.825, respectively. Also, as required under 37 C.F.R. § 1.821(d), amendments to the specification directing reference to the sequence by use of the sequence identifier "SEQ ID NO:" in the text of the description have been provided by way of this amendment under 37 C.F.R. § 1.121. Applicants assert that no new matter has been added. Applicants assert that the amendments made herein are directed to typographical errors, or amendments of the specification to clearly identify sequences referenced in the text of the description to the corrected Sequence Listing, filed herewith.

CONCLUSION

Applicants have amended the specification to clarify the sequences referenced in the specification to the sequences given in the Sequence Listing. Therefore, these amendments do not narrow the scope of the claims within the meaning of *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., Ltd.*, 234 F.3d 558, 586, 56 USPQ2d 1865, 1886 (Fed. Cir. 2000).

Dated: April 4, 2002

Respectfully submitted,

By Elizabeth A. Hart
Elizabeth A. Hart, Registered Patent Agent
Registration No.: 50,931
FULBRIGHT & JAWORSKI L.L.P.
1301 McKinney, Suite 5100
Houston, Texas 77010-3095
(713) 651-5698

Appendix A

Version With Markings to Show Changes Made to Specification

Paragraph on page 6, at lines 2-7:

Typically (i) is any allelic variant of P450_{cam} of *Pseudomonas putida* (e.g. of the polypeptide sequence shown in SEQ ID No. 2 [1]). Typically (ii) is a species homologue of (i) which has sequence homology with (i), and is typically P450_{BM-3} of *Bacillus megaterium* (e.g. the polypeptide sequence shown in SEQ ID No. 4 [2] and the nucleotide sequence is shown in SEQ ID NO: 3), P450_{terp} of *Pseudomonas sp.*, P450_{eryF} of *Saccharopolyspora erythraea*, or P450 105 D1 (CYP105) of *Streptomyces griseus* strains.

Paragraph on page 6, at lines 22-32:

The discussion below provides examples of the positions at which substitutions may be made in P450_{cam}. The same substitutions may be made at equivalent positions in the homologues. Standard nomenclature is used to denote the mutations. The letter of the amino acid present in the natural form is followed by the position, followed by the amino acid in the mutant. To denote multiple mutations in the same protein each mutation is listed separated by hyphens. The mutations discussed below using this nomenclature specify the natural amino acid in P450_{cam}, but it is to be understood that the mutation could be made to a homologue which has a different amino acid at the equivalent position. Note that the amino acid numbering shown in SEQ ID No. 2 [1] for P450_{cam} does not correspond to the numbering used in the description to denote mutations. The numbering in SEQ ID No. 2 [1] is one more than the numbering in the description for a particular position.

Paragraph on page 9, at lines 4-10:

Alternatively a linker may be present between the components. The linker generally comprises amino acids that do not have bulky side chains and therefore do not obstruct the folding of the protein subunits. Preferably the amino acids in the linker are uncharged. Preferred amino acids in the linker are glycine, serine, alanine or threonine. In one embodiment the linker comprises the sequence N-Thr-Asp-Gly-Gly-Ser-Ser-Ser-C (SEQ ID NO:6). The linker is typically from at least 5 amino acids long, such as at least 10, 30 or 50 or more amino acids long.

Paragraph on page 11, at lines 2-6:

In one embodiment the three genes encoding the three proteins [protiens] of the P450_{cam} system, i.e. *camA*, *camB*, and *camC* are placed in the mobile regions of standard transposon vectors and incorporated into the genome of *Pseudomonas* and *flavobacteria*. Alternatively plasmid vectors for expressing these genes may used, in which case the P450_{cam} gene cluster will be extra-chromosomal.

Paragraph on page 14, at lines 2-18:

Oligonucleotide-directed site-specific mutagenesis was carried out by the Kunkel method (Kunkel, T. A. *Proc. Natl. Acad. Sci. USA* 1985, **82**, 488-492) using the Bio-Rad Mutagen kit. The recommended procedure is summarised as follows. An M13 mp19 subclone of the *camC* gene encoding P450_{cam} (SEQ ID NO: 1) was propagated in the *E. coli* strain CJ236. This strain has the *dut*⁻*ung*⁻ phenotype and thus will tolerate the inclusion of uracil in place of thymine in DNA molecules. After three cycles of infection, uracil-containing single stranded (USS) M13 DNA was readily isolated by phenol extraction of mature M13 phage particles excreted into the growth medium. The mutagenic oligonucleotide (or oligonucleotides) were phosphorylated with T4 polynucleotide kinase and then annealed to the USS template. The four nucleotides, DNA polymerase, DNA ligase, ATP and other chemical components were added and the second strand was synthesised *in vitro*. The double stranded form thus obtained was transformed into the *dut*⁺ *ung*⁺ *E. coli* strain MV1190, which should degrade the uracil-containing template strand and propagate the mutant strand synthesised *in vitro*. Plaques were picked and phages of possible mutants grown in *E. coli* strains MV1190 or TG1. The single-stranded DNA from these were sequenced to determine whether the mutagenesis reaction was successful. The mutagenic efficiency was 50 - 80%.

Paragraph on page 16, at lines 19-22:

For the *camA* gene the primer below (SEQ ID NO: 5) was used at the 5' end of the gene to introduce the *Eco* RI cloning site and to change the first codon from GTG to the strong start codon ATG.

5'- GAG ATT AAG AAT TCA TAA ACA CAT GGG AGT GCG TGC CAT ATG AAC
GCA AAC

Eco RI RBS |→*camA*

Paragraph bridging page 16, line 24 to page 17, line 3:

At the 3' end of *camA* the primer (nucleotide sequence is SEQ ID NO: 7) was designed such that 15 bases are complementary to nucleotide sequence of the last five amino acid residues of *camA*. The stop codon immediately

after the GCC codon for the last amino acid was removed, and then part of a seven amino acid linker (Thr Asp Gly Gly Ser Ser Ser; SEQ ID NO: 6) which contained a *Bam* HI cloning site (GGATCC = Gly Ser) was introduced. The coding sequence was thus (amino acid sequence is SEQ ID NO:8):

5'- GAA CTG AGT AGT GCC ACT GAC GGA **GGA TCC** TCA TCG-3'
camA → Thr Asp Gly Gly Ser
|*Bam* HI|

Paragraph on page 17, at lines 5-6:

The primer sequence shown below (SEQ ID NO: 9) is the reverse complement used for PCR:

5'- CGA TGA **GGA TCC** TCC GTC AGT GGC ACT ACT CAG TTC-3'

Paragraph on page 17, at lines 8-13:

For the *camB* gene the primer at the 5' end (nucleotide sequence is SEQ ID NO: 10; amino acid sequence is SEQ ID NO: 11) incorporated the second half of the peptide linker between the reductase and redoxin proteins, and the restriction site *Bam* HI for joining the two amplified genes together.

5'- TCA TCG **GGA TCC** TCA TCG **ATG TCT** AAA GTA GTG TAT-3'
Gly Ser Ser Ser |→ *camB*
|*Bam* HI| Start

Paragraph on page 17, at lines 14-19:

At the 3' end of *camB* the primer incorporates 12 nucleotides complementary to the end of *camB* followed by the stop codon TAA, a 6 nucleotide spacer before the GGAG ribosome binding site. *Xba* I and *Hind* III sites were then added to allow cloning of the *camC* gene when required. The sequence of the coding strand was therefore (SEQ ID NO: 12):

5'- CCC GAT AGG CAA TGG **TAA** TCA TCG **GGAG TCT AGA** GCA TCG **AAG CTT**
TCA TCG-3'
CamB → |stop RBS *Xba* I *Hind* III

Paragraph on page 17, at lines 20-21:

The primer shown below (SEQ ID NO: 13) is the reverse complement used for PCR:

5'-CGA TGA AAG CTT CGA TGC TCT AGA CTCC CGA TGA TTA CCA
TTG CCT ATC GGG -3'

Paragraph on page 19, at lines 2-6:

We used the internal and unique restriction site *Mlu* I (recognition sequence ACGCGT) within the *camB* gene as the starting point so that the PCR product has a different size from the PCR template fragment. The primers were as follows:

5'- TCA TCG **ACG CGT** CGC GAA CTG CTG-3' (SEQ ID NO: 14)

where the *Mlu* I site is in bold.

Paragraph on page 19, at lines 7-10:

The desired coding sequence at the 3' end of the *camB* gene (SEQ ID NO:15) was:

5'- CCC GAT AGG CAA TGG **TAA** GTA GGT GAA TAT CTA ATC CCC ATC
camB → |stop

TAT GCG CGA GTG **GAG TCT** AGA GTT CGA-3'
RBS *Xba* I

Paragraph bridging page 20, lines 25 to page 21, line 7:

The primers shown below maintain the *Hind* III cloning site AAGCTT:

5'-TCA TCG **AAG CTT** GGC TGT TTT-3' (SEQ ID NO: 16)
Hind III |→ vector

At the other end the coding sequence desired was (SEQ ID NO: 17):

5'-ACA ATT TCA CAC **AGGA TCT** AGA C **CAT ATG** TCA TCG **AAG CTT**
TCA TCG-3'
Vector → |RBS *Xba* I *Nde* I *Hind* III

This sequence maintained the *Nde* I and *Hind* III sites but the new *Xba* I site was introduced upstream of the *Nde* I site. The PCR primer used was the reverse complement of the desired sequence (SEQ ID NO: 18):

5'-CGA TGA AAG CTT CGA TGA CAT ATG GTC T AGA TCCT GTG
TGA AAT TGT-3'

Appendix B

Clean Version of Amendments to Specification on April 4, 2002

Paragraph on page 6, at lines 2-7:

Typically (i) is any allelic variant of P450_{cam} of *Pseudomonas putida* (e.g. of the polypeptide sequence shown in SEQ ID No. 2). Typically (ii) is a species homologue of (i) which has sequence homology with (i), and is typically P450_{BM-3} of *Bacillus megaterium* (e.g. the polypeptide sequence shown in SEQ ID No. 4 and the nucleotide sequence is shown in SEQ ID NO: 3), P450_{terp} of *Pseudomonas sp.*, P450_{eryF} of *Saccharopolyspora erythraea*, or P450 105 D1 (CYP105) of *Streptomyces griseus* strains.

Paragraph on page 6, at lines 22-32:

The discussion below provides examples of the positions at which substitutions may be made in P450_{cam}. The same substitutions may be made at equivalent positions in the homologues. Standard nomenclature is used to denote the mutations. The letter of the amino acid present in the natural form is followed by the position, followed by the amino acid in the mutant. To denote multiple mutations in the same protein each mutation is listed separated by hyphens. The mutations discussed below using this nomenclature specify the natural amino acid in P450_{cam}, but it is to be understood that the mutation could be made to a homologue which has a different amino acid at the equivalent position. Note that the amino acid numbering shown in SEQ ID No. 2 for P450_{cam} does not correspond to the numbering used in the description to denote mutations. The numbering in SEQ ID No. 2 is one more than the numbering in the description for a particular position.

Paragraph on page 9, at lines 4-10:

Alternatively a linker may be present between the components. The linker generally comprises amino acids that do not have bulky side chains and therefore do not obstruct the folding of the protein subunits. Preferably the amino acids in the linker are uncharged. Preferred amino acids in the linker are glycine, serine, alanine or threonine. In one embodiment the linker comprises the sequence N-Thr-Asp-Gly-Gly-Ser-Ser-Ser-C (SEQ ID NO:6). The linker is typically from at least 5 amino acids long, such as at least 10, 30 or 50 or more amino acids long.

Paragraph on page 11, at lines 2-6:

In one embodiment the three genes encoding the three proteins of the P450_{cam} system, i.e. camA, camB, and camC are placed in the mobile regions

of standard transposon vectors and incorporated into the genome of *Pseudomonas* and flavobacteria. Alternatively plasmid vectors for expressing these genes may be used, in which case the P450_{cam} gene cluster will be extra-chromosomal.

Paragraph on page 14, at lines 2-18:

Oligonucleotide-directed site-specific mutagenesis was carried out by the Kunkel method (Kunkel, T. A. *Proc. Natl. Acad. Sci. USA* 1985, **82**, 488-492) using the Bio-Rad Mutagen kit. The recommended procedure is summarised as follows. An M13 mp19 subclone of the *camC* gene encoding P450_{cam} (SEQ ID NO: 1) was propagated in the *E. coli* strain CJ236. This strain has the *dut*⁻*ung*⁻ phenotype and thus will tolerate the inclusion of uracil in place of thymine in DNA molecules. After three cycles of infection, uracil-containing single stranded (USS) M13 DNA was readily isolated by phenol extraction of mature M13 phage particles excreted into the growth medium. The mutagenic oligonucleotide (or oligonucleotides) were phosphorylated with T4 polynucleotide kinase and then annealed to the USS template. The four nucleotides, DNA polymerase, DNA ligase, ATP and other chemical components were added and the second strand was synthesised *in vitro*. The double stranded form thus obtained was transformed into the *dut*⁺*ung*⁺ *E. coli* strain MV1190, which should degrade the uracil-containing template strand and propagate the mutant strand synthesised *in vitro*. Plaques were picked and phages of possible mutants grown in *E. coli* strains MV1190 or TG1. The single-stranded DNA from these were sequenced to determine whether the mutagenesis reaction was successful. The mutagenic efficiency was 50 - 80%.

Paragraph on page 16, at lines 19-22:

For the *camA* gene the primer below (SEQ ID NO: 5) was used at the 5' end of the gene to introduce the *Eco* RI cloning site and to change the first codon from GTG to the strong start codon ATG.

5'- GAG ATT AAG AAT TCA TAA ACA CAT GGG AGT GCG TGC CAT ATG AAC
GCA AAC

Eco RI RBS |→*camA*

Paragraph bridging page 16, line 24 to page 17, line 3:

At the 3' end of *camA* the primer (nucleotide sequence is SEQ ID NO: 7) was designed such that 15 bases are complementary to nucleotide sequence of the last five amino acid residues of *camA*. The stop codon immediately after the GCC codon for the last amino acid was removed, and then part of a seven amino acid linker (Thr Asp Gly Gly Ser Ser Ser; SEQ ID NO: 6) which

contained a *Bam* HI cloning site (GGATCC = Gly Ser) was introduced. The coding sequence was thus (amino acid sequence is SEQ ID NO:8):

5'- GAA CTG AGT AGT GCC ACT GAC GGA **GGA TCC** TCA TCG-3'
camA → Thr Asp Gly Gly Ser
|*Bam* HI|

Paragraph on page 17, at lines 5-6:

The primer sequence shown below (SEQ ID NO: 9) is the reverse complement used for PCR:

5'- CGA TGA **GGA TCC** TCC GTC AGT GGC ACT ACT CAG TTC-3'

Paragraph on page 17, at lines 8-13:

For the *camB* gene the primer at the 5' end (nucleotide sequence is SEQ ID NO: 10; amino acid sequence is SEQ ID NO: 11) incorporated the second half of the peptide linker between the reductase and redoxin proteins, and the restriction site *Bam* HI for joining the two amplified genes together.

5'- TCA TCG **GGA TCC** TCA TCG **ATG TCT** AAA GTA GTG TAT-3'
Gly Ser Ser Ser |→ *camB*
|*Bam* HI| Start

Paragraph on page 17, at lines 14-19:

At the 3' end of *camB* the primer incorporates 12 nucleotides complementary to the end of *camB* followed by the stop codon TAA, a 6 nucleotide spacer before the GGAG ribosome binding site. *Xba* I and *Hind* III sites were then added to allow cloning of the *camC* gene when required. The sequence of the coding strand was therefore (SEQ ID NO: 12):

5'- CCC GAT AGG CAA TGG **TAA** TCA TCG **GGAG TCT** AGA GCA TCG **AAG CTT**
TCA TCG-3'
CamB →|stop RBS *Xba* I *Hind* III

Paragraph on page 17, at lines 20-21:

The primer shown below (SEQ ID NO: 13) is the reverse complement used for PCR:

5'-CGA TGA AAG CTT CGA TGC TCT AGA CTCC CGA TGA TTA CCA
TTG CCT ATC GGG -3'

Paragraph on page 19, at lines 2-6:

We used the internal and unique restriction site *Mlu* I (recognition sequence ACGCGT) within the *camB* gene as the starting point so that the

Application No. 10/018,730

Docket No. HO-P02353US1

PCR product has a different size from the PCR template fragment. The primers were as follows:

5'- TCA TCG **ACG CGT** CGC GAA CTG CTG-3' (SEQ ID NO: 14)

where the *Mlu* I site is in bold.

Paragraph on page 19, at lines 7-10:

The desired coding sequence at the 3' end of the *camB* gene (SEQ ID NO:15) was:

5'- CCC GAT AGG CAA TGG **TAA** GTA GGT GAA TAT CTA ATC CCC ATC
camB →|stop

TAT GCG CGA GTG **GAG TCT AGA** GTT CGA-3'
RBS *Xba* I

Paragraph bridging page 20, lines 25 to page 21, line 7:

The primers shown below maintain the *Hind* III cloning site AAGCTT:

5'-TCA TCG **AAG CTT** GGC TGT TTT-3' (SEQ ID NO: 16)
Hind III |→ vector

At the other end the coding sequence desired was (SEQ ID NO: 17):

5'-ACA ATT TCA CAC **AGGA TCT AGA C CAT ATG** TCA TCG **AAG CTT**
TCA TCG-3'
Vector →|RBS *Xba* I *Nde* I *Hind* III

This sequence maintained the *Nde* I and *Hind* III sites but the new *Xba* I site was introduced upstream of the *Nde* I site. The PCR primer used was the reverse complement of the desired sequence (SEQ ID NO: 18):

5'-CGA TGA AAG CTT CGA TGA CAT ATG GTC T AGA TCCT GTG
TGA AAT TGT-3'

PROCESS FOR OXIDISING AROMATIC COMPOUNDS

The invention relates to a process for enzymatically oxidising halogenated aromatic compounds.

Chlorinated aromatic compounds such as the chlorobenzene and polychlorinated biphenyls (PCBs) are among the most wide-spread organic contaminants in the environment due to their common application as solvents, biocides, and in the heavy electrical industry. They are also some of the most problematic environmental pollutant, not only because of the health hazards (lipid solubility and hence accumulation in fatty tissues, toxicity and carcinogenicity) but also because of their slow degradation in the environment.

Whilst microorganisms have shown extraordinary abilities to adapt and evolve to degrade most of the organic chemicals released into the environment, the most chemically inert compounds such as PCBs do persist for two main reasons. First, these compounds have very low solubility in water and therefore their bioavailability is low. Research into this problem has focussed on the use of detergents and other surfactants to enhance their solubility and bioavailability. Second, these compounds require activation by enzymatic oxidation or reduction, and it can take a long time for the necessary genetic adaptations by microorganisms to occur, and even then the organisms may not be stable and viable.

We have now found, according to the present invention, that a monooxygenase, in particular P450_{cam} and its physiological electron transfer partners putidaretoxin and putidaretoxin reductase, can be used to oxidise halogenated aromatic compounds. Also mutants of the monooxygenase with substitutions in the active site have enhanced oxidation activity. Thus suitable monooxygenases can be expressed in microorganisms, animals and plants which are going to be used to oxidise the halogenated aromatic compounds.

Accordingly the present invention provides a process for oxidising a substrate which is a halo aromatic compound, which process comprises oxidising said substrate with a monooxygenase enzyme.

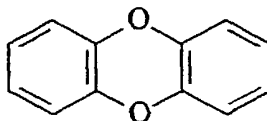
The process may be carried out in a cell that expresses:

- (a) the enzyme
- (b) an electron transfer reductase; and
- (c) an electron transfer redoxin.

The halo aromatic compound is typically a benzene or biphenyl compound. The benzene ring is optionally fused and can be substituted. The halogen is typically chlorine. In many cases there is more than one halogen atom in the molecule, typically 2 to 5 or 6, for example 3. Generally 2 of the halogen atoms will be ortho or para to one another. The compound may or may not contain an oxygen atom such as a hydroxy group, an aryloxy group or a carboxy group. The compound may or may not be chlorophenol or a chlorophenoxyacetic compound.

Specific compounds which can be oxidised by the process of the present invention include 1, 2; 1,3- and 1,4-dichlorobenzene, 1, 2, 4; 1, 2, 3- and 1, 3, 5-trichlorobenzene, 1, 2, 4, 5- and 1, 2, 3, 5- tetrachlorobenzene, pentachlorobenzene, hexachlorobenzene, 3,3'-dichlorobiphenyl and 2, 3, 4, 5, 6- and 2, 2', 4, 5, 5'-pentachlorobiphenyl.

Other compounds which can be oxidised by the process include recalcitrant halo aromatic compounds, especially dioxins and halogenated dibenzofurans, and the corresponding compounds where one or both oxygen atoms is/are replaced by sulphur, in particular compounds of the formula:



which possess at least one halo substituent, such as dioxin itself, 2,3,7,8-tetrachlorodibenzodioxin.

The oxidation typically gives rise to 1, 2 or more oxidation products. These oxidation products will generally comprise 1 or more hydroxyl groups. Generally, therefore, the oxidation products are phenols which can readily be degraded. It is particularly noteworthy that pentachlorobenzene and hexachlorobenzene can be oxidised in this way since they are very difficult to degrade. In contrast the corresponding phenols can be readily degraded by a variety of Pseudomonas and other bacteria. The atom which is oxidized is generally a ring carbon.

The enzyme is typically a natural monooxygenase or a mutant thereof. The

natural monooxygenase is generally a prokaryotic or eukaryotic enzyme. Typically it is a haem-containing enzyme and/or a P450 enzyme. The monooxygenase may or may not be a TfdA (2,4-dichlorophenoxy) acetate/ α -KG dioxygenase. The monooxygenase is generally of microorganism (e.g. bacterial), fungal, yeast, plant or animal origin, typically of a bacterium of the genus *Pseudomonas*. These organisms are typically soil, fresh water or salt water dwelling. In the case of a mutant monooxygenase the non-mutant form may or may not be able to oxidize the substrate.

The monooxygenase typically has a coupling efficiency of at least 1%, such as at least 2%, 4%, 6% or more. The monooxygenase typically has a product formation rate of at least 5 min⁻¹, such as at least 8, 10, 15, 20, 25, 50, 100, 150 min⁻¹ or more. The coupling efficiency or product formation rate is typically measured using any of the substrates or conditions mentioned herein. Thus they are typically measured in the in vitro conditions described in Example 2, in which case the relevant monooxygenase, reductase and redoxin would be present instead of, but at the same concentration as, P450_{cam}, putidaretoxin reductase and putidaretoxin.

The mutant typically has at least one mutation in the active site. A preferred mutant comprises a substitution of an amino acid in the active site by an amino acid with a less polar side chain. Thus the amino acid is typically substituted with an amino acid which is above it in Table 1.

Table 1. HYDROPATHY SCALE FOR AMINO ACID SIDE CHAINS

	Side Chain	Hydropathy
5		
	Ile	4.5
	Val	4.2
10	Leu	3.8
	Phe	2.8
	Cys	2.5
	Met	1.9
	Ala	1.8
15	Gly	-0.4
	Thr	-0.7
	Ser	-0.8
	Trp	-0.9
	Tyr	-1.3
20	Pro	-1.6
	His	-3.2
	Glu	-3.5
	Gln	-3.5
	Asp	-3.5
25	Asn	-3.5
	Lys	-3.9
	Arg	-4.5

An amino acid 'in the active site' is one which lines or defines the site in which the substrate is bound during catalysis or one which lines or defines a site through which the substrate must pass before reaching the catalytic site. Therefore such an amino acid typically interacts with the substrate during entry to the catalytic site or during catalysis. Such an interaction typically occurs through an electrostatic interaction (between charged or polar groups), hydrophobic interaction, hydrogen bonding or van der Waals forces.

The amino acids in the active site can be identified by routine methods to those skilled in the art. These methods include labelling studies in which the enzyme is allowed to bind a substrate which modifies ('labels') amino acids which contact the substrate. Alternatively the crystal structure of the enzyme with bound substrate can be obtained in order to deduce the amino acids in the active site.

The monooxygenase typically has 1, 2, 3, 4 or more other mutations, such as substitutions, insertions or deletions. The other mutations may be in the active site or outside the active site. Typically the mutations are in the 'second sphere' residues which affect or contact the position or orientation of one or more of the amino acids in the active site. The insertion is typically at the N and/or C terminal and thus the enzyme may be part of a fusion protein. The deletion typically comprises the deletion of amino acids which are not involved in catalysis, such as those outside the active site. The monooxygenase may thus comprise only those amino acids which are required for oxidation activity.

The other mutations in the active site typically alter the position and/or conformation of the substrate when it is bound in the active site. The mutation may make the site on the substrate which is to be oxidized more accessible to the haem group. Thus the mutation may be a substitution to an amino acid which has a smaller or larger, or more or less polar, side chain.

The other mutations typically increase the stability of the protein, or make it easier to purify the protein. They typically prevent the dimerisation of the protein, typically by removing cysteine residues from the protein (e.g. by substitution of cysteine at position 334 of P450_{cam}, or at an equivalent position in a homologue, preferably to alanine). They typically allow the protein to be prepared in soluble form, for example by the introduction of deletions or a poly-histidine tag, or by mutation of the N-terminal membrane anchoring sequence. The mutations typically inhibit protein oligomerisation, such as oligomerisation arising from contacts between hydrophobic patches on protein surfaces.

Typically the mutant monooxygenase is at least 70% homologous to a natural monooxygenase on the basis of amino acid identity.

Any of the homologous proteins mentioned herein are typically at least 70% homologous to a protein or at least 80 or 90% and more preferably at least 95%, 97% or 99% homologous thereto over at least 20, preferably at least 30, for instance at least 40, 60 or 100 or more contiguous amino acids. The contiguous amino acids may include the active site. This homology may alternatively be measured not over contiguous amino acids or nucleotides but over only the amino acids in the active site.

The monooxygenase is preferably:

(i) P450_{cam},

(ii) a naturally occurring homologue of (i),

(iii) a mutant of (i) or (ii).

Typically (i) is any allelic variant of P450_{cam} of *Pseudomonas putida* (e.g. of the polypeptide sequence shown in SEQ ID No. 1). Typically (ii) is a species homologue of (i) which has sequence homology with (i), and is typically P450_{BM-3} of *Bacillus megaterium* (e.g. the polypeptide sequence shown in SEQ ID No. 2), P450_{terp} of *Pseudomonas sp.*, P450_{eryF} of *Saccharopolyspora erythraea*, or P450 105 D1 (CYP105) of *Streptomyces griseus* strains.

The active site of (ii) or (iii) may be substantially the same as the active site of (i) or any of the mutants of (i) mentioned herein. Thus the site may comprise the same amino acids in substantially the same positions.

Typically in (iii) amino acid 96 of P450_{cam}, or the equivalent amino acid in a homologue, has been changed to an amino acid with a less polar side chain.

The 'equivalent' side chain in the homologue is one at the homologous position. This can be deduced by lining up the P450_{cam} sequence and the sequence of the homologue based on the homology between the two sequences. The PILEUP, BLAST and BESTFIT algorithms can be used to line up the sequences (for example as described in Altschul S. F. (1993) *J Mol Evol* 36:290-300; Altschul, S, F *et al* (1990) *J Mol Biol* 215:403-10 and (Devereux *et al* (1984) *Nucleic Acids Research* 12, p387-395)). These algorithms can also be used to calculate the levels of homology discussed herein (for example on their default settings). The equivalent amino acid will generally be in a similar place in the active site of the homologue as amino acid 96 in P450_{cam}.

The discussion below provides examples of the positions at which substitutions may be made in P450_{cam}. The same substitutions may be made at equivalent positions in the homologues. Standard nomenclature is used to denote the mutations. The letter of the amino acid present in the natural form is followed by the position, followed by the amino acid in the mutant. To denote multiple mutations in the same protein each mutation is listed separated by hyphens. The mutations discussed below using this nomenclature specify the natural amino acid in P450_{cam}, but it is to be understood that the mutation could be made to a homologue which has a different amino acid at the equivalent position. Note that the amino acid numbering shown in SEQ ID No. 1 for P450_{cam} does not correspond to the numbering used in the description to denote mutations. The numbering in SEQ ID No. 1 is one more than the numbering in the description for a particular position.

An additional mutation is typically an amino acid substitution at amino acid 87, 98, 101, 185, 244, 247, 248, 296, 395, 396 or a combination of these, for example as shown in table 2.

The following combinations of substitutions are preferred:

5 (i) Substitution at position 87 to amino acids of different side-chain volume, such as substitutions (typically of F) to A, L, I and W, combined with substitutions at position 96 to amino acids of different side-chain volume such as (typically Y to) A, L, F, and W. These combinations alter the space available in the upper part of the substrate pocket compared to the wild-type enzyme, for example, from Y96W-F87W (little space) to 10 Y96A-F87A (more space), as well as the location of the space, for example from one side in Y96F-F87A to the other in Y96A-F87W.

(ii) Substitution at position 96 to F combined with substitutions at positions 185 and 395. Both T185 and I395 are at the upper part of the substrate pocket, and substitution with A creates more space while substitution with F will reduce the space 15 available and push the substrate close to the haem.

(iii) Substitutions at position 96 to A, L, F, and W combined with substitutions at residues closer to the haem including at 101, 244, 247, 295, 296 and 396 to A, L, F, or W. These combinations will create or reduce space in the region of the different side-chains to offer different binding orientations to substrates of different sizes. For example, 20 the combinations Y96W-L244A and Y96L-V247W will offer very different pockets for the binding of the substrate.

(iv) Triple substitutions at combinations of positions 87, 96, 244, 247, 295, 296, 395 and 396 with combinations of A, L, F, and W. The aim is to vary the size and shape of the hydrophobic substrate binding pocket. For example, the Y96A-F87A-L244A 25 combination creates more space compared to the Y96F-F87W-V396L combination, thus allowing larger substrates to bind to the former while restricting the available binding orientations of smaller substrates in the latter. The combinations Y96F-F87W-V247L and Y96F-F87W-V295I have comparable substrate pocket volumes, but the locations of the space available for substrate binding are very different. The combination Y96F-F87L- 30 V247A has a slightly larger side-chain volume at the 96 position than the combination Y96L-F87L-V247A, but the L side-chain at the 96 position is much more flexible and the substrate binding orientations will be different for the two triple mutants.

(v) The mutants with four or five substitutions were designed with similar principles of manipulating the substrate volume, the different flexibility of various side-chains, and the location of the space available in the substrate pocket for substrate binding so as to effect changes in selectivity of substrate oxidation.

5 Mutations are generally introduced into the enzyme by using methods known in the art, such as site directed mutagenesis of the enzyme, PCR and gene shuffling methods or by the use of multiple mutagenic oligonucleotides in cycles of site-directed mutagenesis. Thus the mutations may be introduced in a directed or random manner. Typically the mutagenesis method produces one or more polynucleotides encoding one or more different
10 mutants. In one embodiment a library of mutant oligonucleotides is produced which can be used to produce a library of mutant enzymes.

The process is typically carried out in the presence of the natural cofactors of the monooxygenase. Thus typically in addition to the enzyme (a) and the substrate the process is carried out in the presence of an electron transfer reductase (b), an electron transfer
15 redoxin (c), cofactor for the enzyme and an oxygen donor. In this system the flow of electrons is generally: cofactor \rightarrow (b) \rightarrow (c) \rightarrow (a).

(b) is generally an electron transfer reductase which is able to mediate the transfer of electrons from the cofactor to (c), such as a naturally occurring reductase or a protein which has homology with a naturally occurring reductase, such as at least 70%
20 homology; or a fragment of the reductase or homologue. (b) is typically a reductase of any of the organisms mentioned herein, and is typically a flavin dependent reductase, such as putidaredoxin reductase.

(c) is generally an electron transfer redoxin which is able to mediate the transfer of electrons from the cofactor to (a) via (b). (c) is typically a naturally occurring electron
25 transfer redoxin or a protein which has homology with a naturally occurring electron transfer redoxin, such as at least 70% homology; or a fragment of the redoxin or homologue. (c) is typically a redoxin of any of the organisms mentioned herein. (c) is typically a two-iron/two sulphur redoxin, such as putidaredoxin.

The cofactor is any compound capable of donating an electron to (b), such as
30 NADH. The oxygen donor is any compound capable of donating oxygen to (a), such as dioxygen.

Typically (a), (b) and (c) are present as separate proteins; however they may be

water (e.g. fresh water or sea water). When it carried out in culture the culture typically comprises different types of cells of the invention, for example expressing different monooxygenases of the invention. Generally such cells are cultured in the presence of assimilable carbon and nitrogen sources.

5 Typically the cell in which the process is carried out is one in which the monooxygenase does not naturally occur. In another embodiment the monooxygenase is expressed in a cell in which it does naturally occur, but at higher levels than naturally occurring levels. The cell may produce 1, 2, 3, 4 or more different monooxygenases of the invention. These monooxygenases may be capable of oxidising different halo aromatic
10 compounds. Typically the cell also expresses any of the reductases and/or redoxins discussed above.

The cell is typically produced by introducing into a cell (i.e. transforming the cell with) a vector comprising a polynucleotide that encodes the monooxygenase. The vector may integrate into the genome of the cell or remain extrachromosomal. The cell may
15 develop into the animal or plant discussed below. Typically the coding sequence of the polynucleotide is operably linked to a control sequence which is capable of providing for the expression of the coding sequence by the host cell. The control sequence is generally a promoter, typically of the cell in which the monooxygenase expressed.

The term "operably linked" refers to a juxtaposition wherein the components
20 described are in a relationship permitting them to function in their intended manner. A control sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences.

The vector is typically a transposon, plasmid, virus or phage vector. It typically
25 comprises an origin of replication. It typically comprises one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid. The vector is typically introduced into host cells using conventional techniques including calcium phosphate precipitation, DEAE-dextran transfection, or electroporation.

Components (b) and (c) may be expressed in the cell in a similar manner.
30 Typically (a), (b) and (c) are expressed from the same vector, or may be expressed from different vectors. They may be expressed as three different polypeptides. Alternatively they may be expressed in the form of fusion proteins. The cell typically expresses more

than one type of monooxygenase.

In one embodiment the three genes encoding the three proteins of the P450_{cam} system, i.e. camA, camB, and camC are placed in the mobile regions of standard transposon vectors and incorporated into the genome of Pseudomonas and flavobacteria.

5 Alternatively plasmid vectors for expressing these genes may be used, in which case the P450_{cam} gene cluster will be extra-chromosomal.

The cell may be prokaryotic or eukaryotic and is generally any of the cells or of any of the organisms mentioned herein. Preferred cells are Pseudomonas, flavobacteria or fungi cells (e.g. Aspergillus). In one embodiment the cell is one which in its naturally
10 occurring form is able to oxidise any of the substrates mentioned herein. Typically the cell is in a substantially isolated form and/or substantially purified form, in which case it will generally comprise at least 90%, e.g. at least 95%, 98% or 99% of the cells or dry mass of the preparation.

The invention provides a transgenic animal or plant whose cells are any of the
15 cells of the invention. The animal or plant is transgenic for the monooxygenase gene and typically also an appropriate electron transfer reductase and/or redoxin gene. They may be homozygote or heterozygote for such genes, which are typically transiently introduced into the cells, or stably integrated (e.g. in the genome). The animal is typically a worm (e.g. earthworm) or nematode. The plant or animal may be obtained by transforming an
20 appropriate cell (e.g. embryo stem cell, callus or germ cell), fertilising the cell if required, allowing the cell to develop into the animal or plant and breeding the animal or plant true if required. The animal or plant may be obtained by sexual or asexual (e.g. cloning) propagation of an animal or plant of the invention or of the F1 organism (or any generation removed from the F1, or the chimera that develops from the transformed cell).

25 As discussed above the process may be carried out at a locus. Thus the invention also provides a method of treating a locus contaminated with a halo aromatic compound comprising contacting the locus with a monooxygenase, cell, animal or plant of the invention. These organisms are then typically allowed to oxidise the halo aromatic compound. In one embodiment the organisms used to treat the locus are native to the
30 locus. Thus they may be obtained from the locus (e.g. after contamination), transformed/transfected (as discussed above) to express the monooxygenase (and optionally an appropriate electron transfer reductase and/or redoxin).

In one embodiment the locus is treated with more than one type of organism of the invention, e.g. with 2, 3, 4, or more types which express different monooxygenases which oxidise different halo aromatic compounds. In one embodiment such a collection of organisms between them is able to oxidise all halobenzenes, e.g. all chlorobenzenes.

5 The organisms (e.g. in the form of the collection) may carry out the process of the invention in a bioreactor (e.g. in which they are present in immobilised form). Thus the water or soil to be treated may be passed through such a bioreactor. Soil may be washed with water augmented with surfactants or ethanol and then introduced into the bioreactor.

10 The invention also provides a process for selecting a mutant of a monooxygenase for its ability to oxidise any of the substrates mentioned herein, which process comprises screening a library of said mutants for their oxidation effect on the substrate. Thus typically the substrate is provided to the library and mutants are selected based on their ability to oxidise the substrate, for example at a particular rate or under particular conditions. The mutant may be selected based on its ability to oxidise the substrate to a
15 particular oxidation product.

Typically the library will be in the form of cells which comprise the mutant enzymes. Generally each cell will express only one particular mutant enzyme. The library typically comprises at least 500 mutants, such as at least 1,000 or 5,000 mutants, preferably at least 10,000 different mutants.

20 The library typically comprises a random population of mutants. The library may undergo one or more rounds of selection whilst being produced and therefore may not comprise a random population.

25 The library is typically produced by contacting any of the cells discussed herein which expresses the monooxygenase with a mutagen and/or when the cell is a mutator cell culturing the cell in conditions in which mutants are produced. The mutagen may be contacted with the cell prior to or during culturing of the cell. Thus the mutagen may be present during replication of the cell or replication of the genome of the cell.

30 The mutagen generally causes random mutations in the polynucleotide sequence which encodes (a). The mutagen is typically a chemical mutagen, such as nitrosomethylguanidine, methyl- or ethylmethane sulphonic acid, nitrite, hydroxylamine, DNA base analogues, and acridine dyes, such as proflavin. It is typically electromagnetic radiation, such as ultra-violet radiation at 260 nm (absorption maximum of DNA) and X-

rays. It is typically ionising radiation.

A mutator cell is generally deficient in one or more of the primary DNA repair pathways (such as *E. Coli* pathways mutS, mutD or mutT, or their equivalents in another organism), and thus has a high mutation rate. Simply culturing such cell leads to the DNA encoding (a) to become mutated. The cell may be of *E. Coli* XL1 Red mutator strain.

The mutant selected from the library may be used in any aspect of the invention, thus it may be used to oxidise a substrate in the process of the invention or may be expressed in the cell, animal or plant of the invention. It may be used in the method of treating a locus.

The invention is also illustrated by the Examples:

Example 1

Expression of mutants for *in vitro* work.

The P450_{cam} enzymes were expressed using the vector pRH1091 (Baldwin, J.E., Blackburn, J.M., Heath, R.J., and Sutherland, J.D. *Bioorg. Med. Chem. Letts.*, 1992, 2, 663-668.) which utilised the *trc* promoter (a fusion of the *trp* and *lac* promoters). This vector incorporates a strong ribosome binding site (RBS) and the gene to be expressed is cloned using an *Nde* I site on the 5' end of the gene. We used *Hind* III as the cloning site at the 3' end of the *camC* gene. The procedure for protein expression is as follows: Cells are grown at 30°C until the OD_{600nm} reaches 1.0 – 1.2, the temperature is increased to 37°C and camphor added as a 1 M stock in ethanol to a final concentration of 1 mM. The culture is allowed to incubate at 37°C for another 6 hours. The P450_{cam} protein is expressed to high levels in the cytoplasm and the cells take on a red to orange-red colour.

We have also prepared a variant of pRH1091 (by PCR) which has a extra *Xba* I site between the RBS and the *Nde* I site. This is important because *Nde* I is not unique in M13, and this restriction site is also present in the reductase gene as well as the backbone of the pGLW11 vector used for the *in vivo* system. *Xba* I is unique in the polylinker region of M13, but absent in the genes of all three proteins in the P450_{cam} system and in the expression vectors. It therefore allows the *camC* gene to be moved between the mutagenic and expression vectors.

How the mutants were made.

Oligonucleotide-directed site-specific mutagenesis was carried out by the Kunkel method (Kunkel, T. A. *Proc. Natl. Acad. Sci. USA* 1985, 82, 488-492) using the Bio-Rad Mutagen kit. The recommended procedure is summarised as follows. An M13 mp19 subclone of the *camC* gene encoding P450_{cam} was propagated in the *E. coli* strain CJ236. This strain has the *dut*^{ung} phenotype and thus will tolerate the inclusion of uracil in place of thymine in DNA molecules. After three cycles of infection, uracil-containing single stranded (USS) M13 DNA was readily isolated by phenol extraction of mature M13 phage particles excreted into the growth medium. The mutagenic oligonucleotide (or oligonucleotides) were phosphorylated with T4 polynucleotide kinase and then annealed to the USS template. The four nucleotides, DNA polymerase, DNA ligase, ATP and other chemical components were added and the second strand was synthesised *in vitro*. The double stranded form thus obtained was transformed into the *dut*⁺ *ung*⁺ *E. coli* strain MV1190, which should degrade the uracil-containing template strand and propagate the mutant strand synthesised *in vitro*. Plaques were picked and phages of possible mutants grown in *E. coli* strains MV1190 or TG1. The single-stranded DNA from these were sequenced to determine whether the mutagenesis reaction was successful. The mutagenic efficiency was 50 - 80%.

The mutant *camC* gene is excised from the M13 subclone by restriction digest with *Nde* I and *Hind* III, and the fragment of appropriate size is ligated to the backbone of the expression vector prepared by a similar *Nde* I/*Hind* III digest.

Multiple mutants were prepared either by further mutagenesis, also by the Kunkel method, or where the location of the sites in the sequence permits, simple cloning steps. There are two unique restriction sites within the *camC* gene which are absent from the expression vector. One is *Sph* I which spans residues 121 - 123, and the other is *Sal* I which spans residues 338 and 339. Therefore, all mutations at, for example, residues 87, 96, 98, and 101 are readily combined with mutations at higher number residues by ligating appropriate fragments from restriction digests of mutant *camC* genes with *Nde* I/*Sph* I and *Sph* I/*Hind* III and the backbone fragment from a *Nde* I/*Sph* I digest of the expression vector. Mutations at, for example, 395 and 396 can be similarly incorporated by digests in which *Sph* I is replaced with *Sal* I.

The rationale for introducing the unique *Xba* I site is now clear: many mutants with

multiple mutations were prepared by the cloning procedure above. Without the *Xba* I site it would be impossible to clone the gene for these multiple mutants from the expression vector back into M13 for further rounds of mutagenesis. Of course these problems could be overcome by doing mutagenesis by PCR, for example.

5 Example 2

Substrate oxidation protocol: *in vitro* reactions

	Component	Final concentration
	P450 _{cam} enzyme	1 μ M
	Putidaredoxin	10 μ M
10	Putidaredoxin reductase	1 μ M
	Bovine liver catalase	20 μ g/ml
	KCl	200 mM
	Substrate	Typically 1 mM
	NADH	250 - 400 μ M

- 15 * 50 mM Tris-HCl buffer pH 7.4 is added to make up the volume.
 * Temperature controlled at 30°C, optional.
 * The NADH turnover rate could be determined by monitoring the absorbance at 340 nm with time.
 * Catalase does not catalyse the substrate oxidation reactions but rather it is present
 20 to remove any hydrogen peroxide by-product which could otherwise denature the P450_{cam}.

The method can be increased in scale to, for example, 20 ml total incubation volume to allow purification of sufficient products by HPLC for spectroscopic characterisation. Fresh substrate (1 mM) and NADH (1 - 2 mM) are added periodically, such as every 20 minutes in a total reaction time of, typically, 3 hours.

25 Example 3

The *in vivo* system

The *in vivo* systems were expressed using the vector pGLW11, a derivative of the plasmid pKK223 (Brosius, J. and Holy, A. *Proc. Natl. Acad. Sci. USA*, 1984, 81,6929-6933). Expression is directed by the *tac* promoter and the vector incorporates a gene

conferring resistance to the antibiotic ampicillin.

Two systems were constructed. The first one expressed the electron transfer proteins putidaredoxin reductase (*camA* gene) and putidaredoxin (*camB* gene) as a fusion protein with a seven amino acid peptide linker, and the P450_{cam} enzyme (*camC* gene) was expressed by the same vector but it was not fused to the electron proteins. The second system expressed the three proteins as separate entities in the *E.Coli* host. Both systems were catalytically competent for substrate oxidation *in vivo*.

The general strategy was as follows. The genes for the three proteins were cloned using *Eco* RI and *Hind* III as flanking sites, with *Eco* RI at the 5' end. For both *in vivo* systems there are restriction sites between the genes, including between the reductase and redoxin genes in the fusion construct. These restriction sites were introduced by PCR, as detailed below. The first task, however, was to carry out a silent mutation to remove the *Hind* III site within the *camA* gene for the reductase. The AAGCTT *Hind* III recognition sequence in the *camA* gene was changed to AAGCCT, which is a silent mutation because GCT and GCC both encode alanine. The gene was completely sequenced to ensure that there were no spurious mutations.

1. The fusion protein system

1.a Manipulation of the *camA* gene by PCR

For the *camA* gene the primer below was used at the 5' end of the gene to introduce the *Eco* RI cloning site and to change the first codon from GTG to the strong start codon ATG.

5'- GAG ATT AAG AAT TCA TAA ACA CAT GGG AGT GCG TGC CAT ATG AAC GCA AAC
 Eco RI RBS | -*camA*

At the 3' end of *camA* the primer was designed such that 15 bases are complementary to nucleotide sequence of the last five amino acid residues of *camA*. The stop codon immediately after the GCC codon for the last amino acid was removed, and then part of a seven amino acid linker (Thr Asp Gly Gly Ser Ser Ser) which contained a *Bam* HI cloning site (GGATCC = Gly Ser) was introduced. The coding sequence was thus:

WO 00/78973

PCT/GB00/02379

-17-

5'- GAA CTG AGT AGT GCC ACT GAC GGA GGA TCC TCA TCG-3'

camA → Thr Asp Gly Gly Ser

|*Bam* HI|

5 The primer sequence shown below is the reverse complement used for PCR:

5'- CGA TGA GGA TCC TCC GTC AGT GGC ACT ACT CAG TTC-3'

1.b Manipulations of the *camB* gene by PCR

For the *camB* gene the primer at the 5' end incorporated the second half of the peptide linker between the reductase and redoxin proteins, and the restriction site *Bam* HI for joining the two amplified genes together.

5'- TCA TCG GGA TCC TCA TCG ATG TCT AAA GTA GTG TAT-3'

Gly Ser Ser Ser |→ *camB*

|*Bam* HI| Start

At the 3' end of *camB* the primer incorporates 12 nucleotides complementary to the end of *camB* followed by the stop codon TAA, a 6 nucleotide spacer before the GGAG ribosome binding site. *Xba* I and *Hind* III sites were then added to allow cloning of the *camC* gene when required. The sequence of the coding strand was therefore:

5'- CCC GAT AGG CAA TGG TAA TCA TCG GGAG TCT AGA GCA TCG AAG CTT TCA TCG-3'

CamB →|stop RBS *Xba* I *Hind* III

20 The primer shown below is the reverse complement used for PCR:

5'-CGA TGA AAG CTT CGA TGC TCT AGA CTCC CGA TGA TTA CCA TTG CCT ATC GGG -3'

1.c Preparation of the full fusion construct

The *camA* and *camB* genes were amplified by the PCR using the primers described above. The new *camA* was digested with *Eco* RI and *Bam* HI, while the new *CamB* was digested with *Bam* HI and *Hind* III. The pGLW11 expression vector was digested with *Eco* RI and *Hind* III. All three were purified by agarose gel electrophoresis and the three gel slices containing the separate fragments were excised from the gel and ligated together, and then transformed into *E.Coli* DH5α. Successful ligation of all the fragments were

confirmed by a series of restriction digestion experiments, especially the presence of the new and unique *Xba* I site. The entire sequence of the insert from the *Eco* RI site to the *Hind* III site was determined to ensure that all the sequences were correct.

5 The new plasmid, named pSGB^F, was transformed into *E. Coli* and expression of the reductase and redoxin proteins was induced by IPTG. When a purified P450_{cam} enzyme was added to the cell-free extract, substrate oxidation was observed for a variety of substrates.

10 When the *camC* gene is cloned into the pSGB^F plasmid using the *Xba* I and *Hind* III restriction sites, the new recombinant plasmid thus generated expresses the reductase and redoxin as a fusion protein and the P450_{cam} enzyme as a separate entity both from the same mRNA molecule. This *in vivo* system is catalytically competent for terpene oxidation in whole cells.

2. The *in vivo* system with the protein expressed separately

2.a The basic strategy

15 The starting point of the preparation of this *in vivo* system was the recombinant plasmid used to express the *camA* gene for putidaredoxin reductase. The *camA* gene was cloned into the pGLW11 plasmid using the *Eco* RI and *Bam* HI restriction sites, with *Eco* RI being at the 5' end of the gene. Conveniently the polylinker region of the pGLW11 vector has a *Hind* III site downstream of the *Bam* HI site. The *camB* gene was therefore
20 manipulated by PCR such that it can be cloned into pGLW11 using the *Bam* HI and *Hind* III sites. This new plasmid expresses the reductase and redoxin as separate proteins.

The *camB* gene was cloned into pUC118 by the *Bam* HI and *Hind* III cloning sites to express putidaredoxin for our general *in vitro* substrate oxidation work. Therefore, the PCR primer at the 3' end of the *camB* gene was designed to introduce a ribosome binding
25 site and the *Xba* I restriction site upstream of the *Hind* III site so that the *camC* gene can be inserted downstream of *camB* using the *Xba* I and *Hind* III sites. Therefore the three genes were cloned without fusion in the pGLW11 expression vector and arranged in the order 5'-*camA-camB-camC*-3', and each gene has its own RBS to initiate protein synthesis.

2.b Manipulations of the *camB* gene

We used the internal and unique restriction site *Mlu* I (recognition sequence ACGCGT) within the *camB* gene as the starting point so that the PCR product has a different size from the PCR template fragment. The primers were as follows:

5 5'- TCA TCG **ACG CGT** CGC GAA CTG CTG-3'

where the *Mlu* I site is in bold.

The desired coding sequence at the 3' end of the *camB* gene was:

5'- CCC GAT AGG CAA TGG TAA GTA GGT GAA TAT CTA ATC CCC ATC
camB -|stop

10 TAT GCG CGA GTG **GAG TCT** AGA GTT CGA-3'
 RBS *Xba* I

After the stop codon there is a 35 base spacer before the RBS which is used to initiate the synthesis of the P450_{cam} enzyme. The *Xba* I cloning site is located within the spacer between the RBS and the start codon (not in this primer) of the *camC* gene. The
 15 PCR primer used was the reverse complement of the sequence above. The PCR was carried out and the amplified fragment of the appropriate size was purified by agarose gel electrophoresis and the gel slice excised.

One extra step was necessary to complete the construction of the new plasmid. The plasmid for the fusion protein *in vivo* system was digested with *Mlu* I and *Hind* III
 20 restriction enzymes, purified by agarose gel electrophoresis, and the gel slice for the small *camB* fragment excised. The pUC118 plasmid for *camB* expression was similarly digested, and the gel slice for the backbone was excised. By ligating the two fragments together we prepared a new pUC118-based plasmid which had an *Xba* I site followed by an *Hind* III site downstream of the stop codon of *camB*. This new plasmid was digested with the *Mlu* I
 25 and *Xba* I enzymes and the backbone was ligated with the new *camB* fragment described above to generate a plasmid with the following arrangement of the key components:

..lac Promoter..Bam HI..camB gene..spacer..RBS..Xba I..Hind III..

2.c Preparation of the *in vivo* system plasmid

Once the modified *camB* with the *Xba* I and *Hind* III restriction sites and appropriate spacers were prepared, the *in vivo* system was constructed by cloning this into the pGLW11-based plasmid used to express the *camA* gene (reductase protein) using the *Bam* HI and *Hind* III sites. The new *in vivo* system vector has the following arrangement of the key components:

..lac Promoter..Eco IRI..RBS..camA gene..spacer..Bam HI..RBS..camB gene..spacer..RBS..Xba I..Hind III..

This new plasmid, named pSGB⁺, was transformed into *E. Coli* and expression of the reductase and redoxin proteins was induced by IPTG. When a purified P450_{cam} enzyme was added to the cell-free extract, substrate oxidation was observed for a variety of substrates.

When the *camC* gene is cloned into this pSGC⁺ plasmid using the *Xba* I and *Hind* III restriction sites, the new recombinant plasmid thus generated will express the three proteins separately, each under the direction of its own RBS but from the same mRNA molecule. Thus constitutes the *in vivo* system used in the vast majority of our terpene oxidation work.

3. Introduction of an *Xba* I site into pRH1091

This is the final step to enable the *camC* gene to be cloned into the *in vivo* systems by the two cloning sites *Xba*I and *Hind* III. The *Xba* I site was added by PCR of the entire pRH1091 plasmid using two primers. The presence of these two sites will also enable cloning of the *camC* gene into M13 since both *Xba* I and *Hind* III are unique in *camC* and M13.

The primers shown below maintain the *Hind* III cloning site AAGCTT:

5'-TCA TCG AAG CTT GGC TGT TTT-3'

Hind III |→ vector

At the other end the coding sequence desired was:

5'-ACA ATT TCA CAC AGGA TCT AGA C CAT ATG TCA TCG AAG CTT TCA TCG-3'
 Vector - |RBS *Xba* I *Nde* I *Hind* III

This sequence maintained the *Nde* I and *Hind* III sites but the new *Xba* I site was
 5 introduced upstream of the *Nde* I site. The PCR primer used was the reverse complement
 of the desired sequence:

5'-CGA TGA AAG CTT CGA TGA CAT ATG GTC T AGA TCCT GTG TGA AAT TGT-3'

The PCR product was then purified by agarose gel electrophoresis, digested with
Hind III and circularised with T4 DNA ligase. Success of the PCR method was indicated
 10 by the presence of a new and unique *Xba* I site in plasmid DNA isolated from
 transformants.

4. Cloning of *camC* into the *in vivo* systems

All existing *camC* mutants were cut out of pRH1091-based expression plasmids with
Nde I and *Hind* III. The new vector is similarly cut with the same restriction enzymes and
 15 the *camC* gene cloned into this plasmid with T4 DNA ligase. This DNA is transformed
 into *E. Coli* JM109 which then may be grown to express P450_{cam}.

The *camC* gene is excised from the new vector using *Xba* I and *Hind* III restriction
 enzymes and cloned into either the *in vivo* vector systems or M13mp19 for mutagenesis.

20 5. *In vivo* expression and substrate turnover

For protein expression, cells are grown in LBamp medium (tryptone 10 g/litre,
 yeast extract 5 g/litre, NaCl 10 g/litre, 50 µg/ml ampicillin) at 30°C until the OD_{600nm}
 reaches 1.0 - 1.2. IPTG (isopropyl-β-D-thiogalactopyranoside) was added to a final
 concentration of 1 µM (from a 1 M stock in H₂O) and the culture was incubated at 30°C
 25 overnight.

For simple screening the substrate can be added to culture and the incubation
 continued. However, due to impurities from the culture media the cells were generally
 washed twice with 0.5 vol. of buffer P, (KH₂PO₄ 6.4 g, K₂HPO₄·3H₂O 25.8 g, H₂O to 4

litres, pH 7.4) and resuspended in 0.25 vol. oxygen saturated buffer P containing 24 mM glucose. Substrate was added to 1 mM and the incubation continued at 30°C. The reaction was allowed to run for 24 hours with periodic additions of substrate and glucose.

Example 4

5 The oxidation of halo aromatic compounds

10

Mutant	2,3,6- Trichlorophenol	3,4,6- Trichlorophenol	Coupling Efficiency (%)	Product formation rate (min ⁻¹)
Y96F	75	25	18	22
Y96A	77	23	14	33
Y96H	54	46	3	1
F87L-Y96F	42	58	4	8
F87A-Y96F	52	48	2	3
F87A-Y96-F-V247A	43	57	4	7

15

Mutant	2,3- Dichlorophenol	3,4- Dichlorophenol	Coupling Efficiency (%)	Product formation rate (min ⁻¹)
Y96A	94	6	6	19
Y96F	91	9	4	8
Y96A-V247L	94	6	7	20
Y96L-V247A	90	10	2	0.7
F87L-96F	96	4	3	5
C334A	95	5	2	0.5

20

All mutants have C334A. Coupling efficiency is the percentage of NADH consumed which was utilised for product formation, i.e. a percentage of the theoretical maximum efficiency. The product formation rates are given in (nmol product) (nmol P450_{cam})⁻¹ (min)⁻¹. The relative amount of product formed in each case is shown.

1,3- and 1,4-dichlorobenzene, 1, 2, 3- and 1, 3, 5-trichlorobenzene, 1, 2, 4, 5- and 1, 2, 3, 5-tetrachlorobenzene, and 2, 3, 4, 5, 6- and 2, 2', 4, 5, 5'- pentachlorobiphenyl were also found to be oxidised.

Wild-type and mutant P450_{cam} enzymes were tested for their ability to oxidise 3,3'-dichlorobiphenyl and 2,2',4,5,5'-pentachlorobiphenyl. Results are shown in terms of NADH turnover. Rates are given as nanomol NADH consumed per nanomol P450_{cam} enzyme per minute.

P450 _{cam} enzyme	3,3'-	2,2',4,5,5'-
	dichlorobiphenyl	pentachlorobiphenyl
Wild-type	0.4	not detected
Y96F	15	1
F87A-Y96F	845	165
F87L-Y96F	174	13
F87W-Y96F	4	3
F87A-Y96F-V247A	112	12
Y96A-V247L	84	37
F87A-Y96F-L244A	669	321
F87A-Y97F-L244A-V247A	173	214

The first product, 4-hydroxy-3,3'-dichlorobiphenyl was identified by the characteristic coupling patterns expected in the ¹H NMR spectrum and by mass spectroscopy. The further oxidation product, 4,4'-dihydroxy-3,3'-dichlorobiphenyl was identified by co-elution with an authentic sample, and by UV-vis and mass spectroscopy. This product did not constitute more than ca. 10% of the total products in any of the mutants tested.

For the second substrate product was established as 4'-hydroxy-2,2',4,5,5'-pentachlorobiphenyl by the observation of the parent ion in the mass spectrum, and by comparison with literature ¹H NMR data.

WO 00/78973

PCT/GB00/02379

-24-

P450_{can} mutants

All mutants optionally contain the base mutation C334A.

Single mutants: Y96A, Y96F, Y96L, Y96W.

Double mutants:

Y96A-F87A	Y96F-F87A	Y96F-V295A	Y96L-F87A	Y96L-A296L
Y96A-F87L	Y96F-F87I	Y96F-V295L	Y96L-F87L	Y96L-A296F
Y96A-F87W	Y96F-F87L	Y96F-V295I	Y96L-F98W	Y96L-V396A
Y96A-F98W	Y96F-F87W	Y96F-A296L	Y96L-T101L	Y96L-V396L
Y96A-L244A	Y96F-F98W	Y96F-A296F	Y96L-T101F	Y96L-V396F
Y96A-V247A	Y96F-T101L	Y96F-I395F	Y96L-L244A	Y96L-V396W
Y96A-V247L	Y96F-T101F	Y96F-I395G	Y96L-L244F	
Y96A-I395F	Y96F-T185A	Y96F-V396A	Y96L-V247A	
Y96A-I395G	Y96F-T185F	Y96F-V396L	Y96L-V247L	Y96W-F87W
	Y96F-T185L	Y96F-V396F	Y96L-V247F	Y96W-F98W
	Y96F-L244A	Y96F-V396W	Y96L-V247W	Y96W-L244A
	Y96F-V247A		Y96L-G248L	Y96W-V247A
	Y96F-V247L		Y96L-V295L	Y96W-V396A
	Y96F-G248L		Y96L-V295F	

Triple Mutants:

Y96A-F87A-L244A	Y96L-V247A-V396L	Y96F-F87W-V247A
Y96A-F87A-V247A	Y96L-V247A-V396F	Y96F-F87W-V247L
Y96A-F87L-L244A	Y96L-V247A-V396W	Y96F-F87W-V247F
Y96A-F87L-V247A	Y96L-V247F-V396A	Y96F-F87W-V295L
Y96A-L244A-V247A		Y96F-F87W-A296L
	Y96F-F87A-L244A	Y96F-F87W-V396A
Y96L-F87A-L244A	Y96F-F87A-V247A	Y96F-F87W-V396L
Y96L-F87A-V247A	Y96F-F87A-V247L	Y96F-V247F-V396A
Y96L-F87L-L244A	Y96F-F87A-I395F	Y96F-L244A-V396L
Y96L-F87L-V247A	Y96F-F87A-I395G	Y96F-L244A-V396F
Y96L-V247A-I395F	Y96F-F87L-V247A	Y96F-L244A-V396W
Y96L-V247L-I395F	Y96F-F87L-V247L	Y96F-L244F-V396A
Y96L-V247L-I395G	Y96F-F87L-I395F	Y96F-V247A-V396L
Y96L-L244A-V396L	Y96F-F87W-T185A	Y96F-V247A-V396F
Y96L-L244A-V396F	Y96F-F87W-T185F	Y96F-V247A-V396W
Y96L-L244A-V396W	Y96F-F87W-T185L	
Y96L-L244F-V396A	Y96F-F87W-L244F	Y96W-F87W-F98W

<u>Four mutations:</u>	<u>Five mutations:</u>
Y96A-F87A-L244A-V247A	Y96F-F87W-T185L-V247L-V295L
Y96A-F87L-L244A-V247A	Y96F-F87W-T185L-V247L-V396A
Y96L-F87A-L244A-V247A	Y96F-F87W-T185L-V247L-V396L
Y96L-F87L-L244A-V247A	
Y96F-F87W-L244A-V295L	
Y96F-F87W-L244F-V396A	
Y96F-F87W-L244A-A296L	
Y96F-F87W-V247A-V396L	
Y96F-F87W-V247A-V396F	
Y96F-F87W-V247L-V295A	
Y96F-F87W-V247L-V396A	
Y96F-F87W-V247F-V396A	
Y96F-F87W-V247A-I395F	
Y96F-F87W-V247L-I395G	

Table 2 (continued)

1. Process for oxidising a substrate which is a halo aromatic compound, which process comprises oxidising said substrate with a monooxygenase enzyme.
2. Process according to claim 1 in which the enzyme comprises a substitution of an amino acid in the active site by an amino acid with a less polar side-chain.
3. Process according to claim 2 in which the enzyme comprises one or more other amino acid substitutions in the active site.
4. Process according to any one of the preceding claims in which the enzyme is
- (i) P450_{cam}, or
- (ii) a naturally occurring homologue of (i), or
- (iii) a mutant of (i) or (ii).
5. Process according to claim 4 in which the enzyme is one in which amino acid 96 of P450_{cam}, or the equivalent amino acid in a homologue, has been changed to an amino acid with a less polar side-chain.
6. Process according to any one of the preceding claims in which the halogen is chlorine.
7. Process according to any one of the preceding claims in which the aromatic compound is a benzene or biphenyl.
8. Process according to any one of the preceding claims in which the substrate has more than one halogen atom.
9. Process according to claim 8 in which the substrate is 1, 2-dichlorobenzene, 1, 2, 4- trichlorobenzene, 3,3'-dichlorobiphenyl or 2,2',4,5,5'-pentachlorobiphenyl.
10. Process according to claim 8 in which the substrate is pentachlorobenzene or hexachlorobenzene.
11. Process according to any one of the preceding claims which is carried out in a cell that expresses:
- (a) an enzyme as defined in any one of claims 1 to 5;
- (b) an electron transfer reductase; and
- (c) an electron transfer redoxin.
12. Process according to claim 11 in which:
- (b) is putidaretoxin reductase or a homologue; or a fragment thereof; and/or

-27-

(c) is putidaretoxin or a homologue; or a fragment thereof.

13. Process according to claim 11 or 12 wherein the cell is one in which the enzyme (a) does not naturally occur.

5 14. Process according to any one of claims 11 to 13 wherein the cell is one which in its naturally occurring form is able to oxidise a substrate as defined in any one of claims 6 to 10.

15. A cell as defined in claim 14.

16. A transgenic animal or plant whose cells are as defined in any one of claims 11 to 14.

10 17. Method of treating a locus contaminated with a substrate as defined in any one of claims 1 or 6 to 10 comprising contacting the locus with an enzyme as defined in any one of claims 1 to 5 or a cell as defined in any of claims 11 to 13, or an animal or plant as defined in claim 16.

15 18. Process for selecting a mutant of an enzyme as defined in claim 1, 4(i) or 4(ii) for its ability to oxidise a substrate as defined in claim 1, or any one of the claims 6 to 10, which process comprises screening a library of said mutants for their oxidation effect on the substrate.

19. Process, cell, animal, plant or method according to any one of claims 1 to 17 wherein the enzyme is one that has been selected in a process according to claim 18.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
28 December 2000 (28.12.2000)

PCT

(10) International Publication Number
WO 00/78973 A1

(51) International Patent Classification: C12N 15/53,
C12P 7/22, C12N 9/02, A01K 67/033, A01H 5/00, A62D
3/00

(74) Agent: ELLIS-JONES, Patrick, George, Armine; J.A.
Kemp & Co., 14 South Square, Gray's Inn, London WC1R
5LX (GB).

(21) International Application Number: PCT/GB00/02379

(22) International Filing Date: 19 June 2000 (19.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
9914373.7 18 June 1999 (18.06.1999) GB

(71) Applicant (for all designated States except US): ISIS IN-
NOVATION LIMITED [GB/GB]; Ewert House, Ewert
Place, Summertown, Oxford OX2 7BZ (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): WONG, Luet, Lok
[GB/GB]; University of Oxford, Dept. of Chemistry, South
Parks Road, Oxford OX1 3QR (GB). JONES, Jonathan,
Peter [GB/GB]; University of Oxford, Dept. of Chemistry,
South Parks Road, Oxford OX1 3QR (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

- With international search report.
- Before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments.

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: PROCESS FOR OXIDISING AROMATIC COMPOUNDS

(57) Abstract: Process for oxidising a substrate which is a halo aromatic compound, which process comprises oxidising said sub-
strate with a monooxygenase enzyme. The enzyme may be P450cam. The process may be carried out in cells, animals or plants.

WO 00/78973 A1

95

WO 00/78973

PCT/GB00/02379

-3-

	atg tgt ggc ctg tta ctg gtc ggc ggc ctg gat acg gtg gtc aat ttc	768
	Met Cys Gly Leu Leu Leu Val Gly Gly Leu Asp Thr Val Val Asn Phe	
	245 250 255	
5	ctc agc ttc agc atg gag ttc ctg gcc aaa agc ccg gag cat cgc cag	816
	Leu Ser Phe Ser Met Glu Phe Leu Ala Lys Ser Pro Glu His Arg Gln	
	260 265 270	
	gag ctg atc gag cgt ccc gag cgt att cca gcc gct tgc gag gaa cta	864
	Glu Leu Ile Glu Arg Pro Glu Arg Ile Pro Ala Ala Cys Glu Glu Leu	
	275 280 285	
10	ctc cgg cgc ttc tcg ctg gtt gcc gat ggc cgc atc ctc acc tcc gat	912
	Leu Arg Arg Phe Ser Leu Val Ala Asp Gly Arg Ile Leu Thr Ser Asp	
	290 295 300	
	tac gag ttt cat ggc gtg caa ctg aag aaa ggt gac cag atc ctg cta	960
	Tyr Glu Phe His Gly Val Gln Leu Lys Lys Gly Asp Gln Ile Leu Leu	
15	305 310 315 320	
	ccg cag atg ctg tct ggc ctg gat gag cgc gaa aac gcc tgc ccg atg	1008
	Pro Gln Met Leu Ser Gly Leu Asp Glu Arg Glu Asn Ala Cys Pro Met	
	325 330 335	
	cac gtc gac ttc agt cgc caa aag gtt tca cac acc acc ttt ggc cac	1056
20	His Val Asp Phe Ser Arg Gln Lys Val Ser His Thr Thr Phe Gly His	
	340 345 350	
	ggc agc cat ctg tgc ctt ggc cag cac ctg gcc cgc cgg gaa atc atc	1104
	Gly Ser His Leu Cys Leu Gly Gln His Leu Ala Arg Arg Glu Ile Ile	
	355 360 365	
25	gtc acc ctc aag gaa tgg ctg acc agg att cct gac ttc tcc att gcc	1152
	Val Thr Leu Lys Glu Trp Leu Thr Arg Ile Pro Asp Phe Ser Ile Ala	
	370 375 380	

gaa tca cgc ttt gat aaa aac tta agt caa gcg ctt aaa ttt gta cgt 240
Glu Ser Arg Phe Asp Lys Asn Leu Ser Gln Ala Leu Lys Phe Val Arg
65 70 75 80

WO 00/78973

PCT/GB00/02379

-5-

	gat ttt gca gga gac ggg tta ttt aca agc tgg acg cat gaa aaa aat	288
	Asp Phe Ala Gly Asp Gly Leu Phe Thr Ser Trp Thr His Glu Lys Asn	
	85 90 95	
5	tgg aaa aaa gcg cat aat atc tta ctt cca agc ttc agt cag cag gca	336
	Trp Lys Lys Ala His Asn Ile Leu Leu Pro Ser Phe Ser Gln Gln Ala	
	100 105 110	
	atg aaa ggc tat cat gcg atg atg gtc gat atc gcc gtg cag ctt gtt	384
	Met Lys Gly Tyr His Ala Met Met Val Asp Ile Ala Val Gln Leu Val	
	115 120 125	
10	caa aag tgg gag cgt cta aat gca gat gag cat att gaa gta ccg gaa	432
	Gln Lys Trp Glu Arg Leu Asn Ala Asp Glu His Ile Glu Val Pro Glu	
	130 135 140	
	gac atg aca cgt tta acg ctt gat aca att ggt ctt tgc ggc ttt aac	480
	Asp Met Thr Arg Leu Thr Leu Asp Thr Ile Gly Leu Cys Gly Phe Asn	
15	145 150 155 160	
	tat cgc ttt aac agc ttt tac cga gat cag cct cat cca ttt att aca	528
	Tyr Arg Phe Asn Ser Phe Tyr Arg Asp Gln Pro His Pro Phe Ile Thr	
	165 170 175	
	agt atg gtc cgt gca ctg gat gaa gca atg aac aag ctg cag cga gca	576
20	Ser Met Val Arg Ala Leu Asp Glu Ala Met Asn Lys Leu Gln Arg Ala	
	180 185 190	
	aat cca gac gac cca gct tat gat gaa aac aag cgc cag ttt caa gaa	624
	Asn Pro Asp Asp Pro Ala Tyr Asp Glu Asn Lys Arg Gln Phe Gln Glu	
	195 200 205	
25	gat atc aag gtg atg aac gac cta gta gat aaa att att gca gat cgc	672
	Asp Ile Lys Val Met Asn Asp Leu Val Asp Lys Ile Ile Ala Asp Arg	
	210 215 220	

	aaa gca agc ggt gaa caa agc gat gat tta tta acg cat atg cta aac	720
	Lys Ala Ser Gly Glu Gln Ser Asp Asp Leu Leu Thr His Met Leu Asn	
	225 230 235 240	
5	gga aaa gat cca gaa acg ggt gag ccg ctt gat gac gag aac att cgc	768
	Gly Lys Asp Pro Glu Thr Gly Glu Pro Leu Asp Asp Glu Asn Ile Arg	
	245 250 255	
	tat caa att att aca ttc tta att gcg gga cac gaa aca aca agt ggt	816
	Tyr Gln Ile Ile Thr Phe Leu Ile Ala Gly His Glu Thr Thr Ser Gly	
	260 265 270	
10	ctt tta tca ttt gcg ctg tat ttc tta gtg aaa aat cca cat gta tta	864
	Leu Leu Ser Phe Ala Leu Tyr Phe Leu Val Lys Asn Pro His Val Leu	
	275 280 285	
	caa aaa gca gca gaa gaa gca gca cga gtt cta gta gat cct gct cca	912
	Gln Lys Ala Ala Glu Glu Ala Ala Arg Val Leu Val Asp Pro Ala Pro	
15	290 295 300	
	agc tac aaa caa gtc aaa cag ctt aaa tat gtc ggc atg gtc tta aac	960
	Ser Tyr Lys Gln Val Lys Gln Leu Lys Tyr Val Gly Met Val Leu Asn	
	305 310 315 320	
	gaa gcg ctg cgc tta tgg cca act gct cct gcg ttt tcc cta tat gca	1008
20	Glu Ala Leu Arg Leu Trp Pro Thr Ala Pro Ala Phe Ser Leu Tyr Ala	
	325 330 335	
	aaa gaa gat acg gtg ctt gga gga gaa tat cct tta gaa aaa ggc gac	1056
	Lys Glu Asp Thr Val Leu Gly Gly Glu Tyr Pro Leu Glu Lys Gly Asp	
	340 345 350	
25	gaa cta atg gtt ctg att cct cag ctt cac cgt gat aaa aca att tgg	1104
	Glu Leu Met Val Leu Ile Pro Gln Leu His Arg Asp Lys Thr Ile Trp	
	355 360 365	

WO 00/78973

PCT/GB00/02379

-8-

	cag gtc gca acg ctt gat tca cac gcc gga aat ctt ccg cgc gaa gga	1584
	Gln Val Ala Thr Leu Asp Ser His Ala Gly Asn Leu Pro Arg Glu Gly	
	515 520 525	
5	gct gta tta att gta acg gcg tct tat aac ggt cat ccg cct gat aac	1632
	Ala Val Leu Ile Val Thr Ala Ser Tyr Asn Gly His Pro Pro Asp Asn	
	530 535 540	
	gca aag caa ttt gtc gac tgg tta gac caa gcg tct gct gat gaa gta	1680
	Ala Lys Gln Phe Val Asp Trp Leu Asp Gln Ala Ser Ala Asp Glu Val	
	545 550 555 560	
10	aaa ggc gtt cgc tac tcc gta ttt gga tgc gcc gat aaa aac tgg gct	1728
	Lys Gly Val Arg Tyr Ser Val Phe Gly Cys Gly Asp Lys Asn Trp Ala	
	565 570 575	
	act acg tat caa aaa gtg cct gct ttt atc gat gaa acg ctt gcc gct	1776
	Thr Thr Tyr Gln Lys Val Pro Ala Phe Ile Asp Glu Thr Leu Ala Ala	
15	580 585 590	
	aaa ggg gca gaa aac atc gct gac cgc ggt gaa gca gat gca agc gac	1824
	Lys Gly Ala Glu Asn Ile Ala Asp Arg Gly Glu Ala Asp Ala Ser Asp	
	595 600 605	
	gac ttt gaa ggc aca tat gaa gaa tgg cgt gaa cat atg tgg agt gac	1872
20	Asp Phe Glu Gly Thr Tyr Glu Glu Trp Arg Glu His Met Trp Ser Asp	
	610 615 620	
	gta gca gcc tac ttt aac ctc gac att gaa aac agt gaa gat aat aaa	1920
	Val Ala Ala Tyr Phe Asn Leu Asp Ile Glu Asn Ser Glu Asp Asn Lys	
	625 630 635 640	
25	tct act ctt tca ctt caa ttt gtc gac agc gcc gcg gat atg ccg ctt	1968
	Ser Thr Leu Ser Leu Gln Phe Val Asp Ser Ala Ala Asp Met Pro Leu	
	645 650 655	

WO 00/78973

PCT/GB00/02379

-9-

	gcg aaa atg cac ggt gcg ttt tca acg aac gtc gta gca agc aaa gaa	2016
	Ala Lys Met His Gly Ala Phe Ser Thr Asn Val Val Ala Ser Lys Glu	
	660 665 670	
5	ctt caa cag cca ggc agt gca cga agc acg cga cat ctt gaa att gaa	2064
	Leu Gln Gln Pro Gly Ser Ala Arg Ser Thr Arg His Leu Glu Ile Glu	
	675 680 685	
	ctt cca aaa gaa gct tct tat caa gaa gga gat cat tta ggt gtt att	2112
	Leu Pro Lys Glu Ala Ser Tyr Gln Glu Gly Asp His Leu Gly Val Ile	
	690 695 700	
10	cct cgc aac tat gaa gga ata gta aac cgt gta aca gca agg ttc ggc	2160
	Pro Arg Asn Tyr Glu Gly Ile Val Asn Arg Val Thr Ala Arg Phe Gly	
	705 710 715 720	
	cta gat gca tca cag caa atc cgt ctg gaa gca gaa gaa gaa aaa tta	2208
	Leu Asp Ala Ser Gln Gln Ile Arg Leu Glu Ala Glu Glu Glu Lys Leu	
15	725 730 735	
	gct cat ttg cca ctc gct aaa aca gta tcc gta gaa gag ctt ctg caa	2256
	Ala His Leu Pro Leu Ala Lys Thr Val Ser Val Glu Glu Leu Leu Gln	
	740 745 750	
	tac gtg gag ctt caa gat cct gtt acg cgc acg cag ctt cgc gca atg	2304
20	Tyr Val Glu Leu Gln Asp Pro Val Thr Arg Thr Gln Leu Arg Ala Met	
	755 760 765	
	gct gct aaa acg gtc tgc ccg ccg cat aaa gta gag ctt gaa gcc ttg	2352
	Ala Ala Lys Thr Val Cys Pro Pro His Lys Val Glu Leu Glu Ala Leu	
	770 775 780	
25	ctt gaa aag caa gcc tac aaa gaa caa gtg ctg gca aaa cgt tta aca	2400
	Leu Glu Lys Gln Ala Tyr Lys Glu Gln Val Leu Ala Lys Arg Leu Thr	
	785 790 795 800	

WO 00/78973

PCT/GB00/02379

-10-

	atg ctt gaa ctg ctt gaa aaa tac ccg gcg tgt gaa atg aaa ttc agc	2448
	Met Leu Glu Leu Leu Glu Lys Tyr Pro Ala Cys Glu Met Lys Phe Ser	
	805 810 815	
5	gaa ttt atc gcc ctt ctg cca agc ata cgc ccg cgc tat tac tcg att	2496
	Glu Phe Ile Ala Leu Leu Pro Ser Ile Arg Pro Arg Tyr Tyr Ser Ile	
	820 825 830	
	tct tca tca cct cgt gtc gat gaa aaa caa gca agc atc acg gtc agc	2544
	Ser Ser Ser Pro Arg Val Asp Glu Lys Gln Ala Ser Ile Thr Val Ser	
	835 840 845	
10	gtt gtc tca gga gaa gcg tgg agc gga tat gga gaa tat aaa gga att	2592
	Val Val Ser Gly Glu Ala Trp Ser Gly Tyr Gly Glu Tyr Lys Gly Ile	
	850 855 860	
	gcg tcg aac tat ctt gcc gag ctg caa gaa gga gat acg att acg tgc	2640
	Ala Ser Asn Tyr Leu Ala Glu Leu Gln Glu Gly Asp Thr Ile Thr Cys	
15	865 870 875 880	
	ttt att tcc aca ccg cag tca gaa ttt acg ctg cca aaa gac cct gaa	2688
	Phe Ile Ser Thr Pro Gln Ser Glu Phe Thr Leu Pro Lys Asp Pro Glu	
	885 890 895	
	acg ccg ctt atc atg gtc gga ccg gga aca ggc gtc gcg ccg ttt aga	2736
20	Thr Pro Leu Ile Met Val Gly Pro Gly Thr Gly Val Ala Pro Phe Arg	
	900 905 910	
	ggc ttt gtg cag gcg cgc aaa cag cta aaa gaa caa gga cag tca ctt	2784
	Gly Phe Val Gln Ala Arg Lys Gln Leu Lys Glu Gln Gly Gln Ser Leu	
	915 920 925	
25	gga gaa gca cat tta tac ttc ggc tgc cgt tca cct cat gaa gac tat	2832
	Gly Glu Ala His Leu Tyr Phe Gly Cys Arg Ser Pro His Glu Asp Tyr	
	930 935 940	

WO 00/78973

PCT/GB00/02379

-11-

	ctg tat caa gaa gag ctt gaa aac gcc caa agc gaa ggc atc att acg	2880
	Leu Tyr Gln Glu Glu Leu Glu Asn Ala Gln Ser Glu Gly Ile Ile Thr	
	945 950 955 960	
5	ctt cat acc gct ttt tct cgc atg cca aat cag ccg aaa aca tac gtt	2928
	Leu His Thr Ala Phe Ser Arg Met Pro Asn Gln Pro Lys Thr Tyr Val	
	965 970 975	
	cag cac gta atg gaa caa gac ggc aag aaa ttg att gaa ctt ctt gat	2976
	Gln His Val Met Glu Gln Asp Gly Lys Lys Leu Ile Glu Leu Leu Asp	
	980 985 990	
10	caa gga gcg cac ttc tat att tgc gga gac gga agc caa atg gca cct	3024
	Gln Gly Ala His Phe Tyr Ile Cys Gly Asp Gly Ser Gln Met Ala Pro	
	995 1000 1005	
	gcc gtt gaa gca acg ctt atg aaa agc tat gct gac gtt cac caa gtg	3072
	Ala Val Glu Ala Thr Leu Met Lys Ser Tyr Ala Asp Val His Gln Val	
15	1010 1015 1020	
	agt gaa gca gac gct cgc tta tgg ctg cag cag cta gaa gaa aaa ggc	3120
	Ser Glu Ala Asp Ala Arg Leu Trp Leu Gln Gln Leu Glu Glu Lys Gly	
	1025 1030 1035 1040	
	cga tac gca aaa gac gtg tgg gct ggg taa	3150
20	Arg Tyr Ala Lys Asp Val Trp Ala Gly	
	1045 1050	

N.76277B

I hereby certify that this correspondence is being deposited with the U.S. Postal Service as Express Mail, Airbill No. EK1027032 US in an envelope address to: Commissioner for Patents, Washington, DC 20231, on the date shown below

Dated: 12/18/01 Signature

Colby G. Delgado
Colby G. Delgado

Docket No. P02353US0
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:

Luet L. Wong et al.

Customer No.: 026271

Application No.: Unassigned

Group Art Unit:

Filed: Herewith

Examiner:

For: PROCESS FOR OXIDISING AROMATIC COMPOUNDS

POWER OF ATTORNEY

Commissioner for Patents
Washington, DC 20231

Dear Sir:

Isis Innovation Limited, assignee of the entire right, title and interest in the above-identified application by assignment hereby appoints the members of the firm of Fulbright & Jaworski L.L.P., a firm composed of:

Melissa W. Acosta	<u>45,872</u>	Paul E. Krieger	<u>25,886</u>	John E. Schneider	<u>31,998</u>
Alberto Q. Amatong	<u>41,580</u>	Michael S. McCoy	<u>46,913</u>	Jan K. Simpson	<u>33,283</u>
Ronald G. Bliss	<u>28,691</u>	John M. Mings	<u>35,955</u>	Melissa L. Sistrunk	<u>45,579</u>
Gino Catena	<u>45,546</u>	Thomas D. Paul	<u>32,714</u>	Edward D. Steakley	<u>47,964</u>
Paul L. Deverter II	<u>19,747</u>	Jayne C. Piana	<u>P-48,424</u>	William A. Stout	<u>18,773</u>
Mark L. Delflache	<u>28,942</u>	Doak C. Procter, IV	<u>P-48,705</u>	Charles B. Walker, Jr.,	<u>43,231</u>
David L. Fox	<u>40,612</u>	James W. Repass	<u>30,487</u>	Richard S. Zembek	<u>43,306</u>
Eric B. Hall	<u>46,751</u>				

as its attorneys with full power of substitution to prosecute this application and to transact all business in the Patent and Trademark Office in connection therewith.

The assignee certifies that it has reviewed the assignment and to the best of the assignee's knowledge and belief, title is in the assignee.

Application No.

Docket No.

Please direct all correspondence regarding this application to the following:

FULBRIGHT & JAWORSKI L.L.P.

Attn: Melissa L. Sistrunk

1301 McKinney, Suite 5100

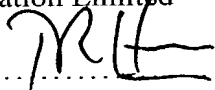
Houston, Texas 77010-3095

Telephone: (713) 651-3735

Fax: (713) 651-5246

For: Isis Innovation Limited

By:



Dated:

3 - DEC - 2001

Printed Name:

T. R. HOCKADAY

N.76277B

Attorney's Docket No.

**RULE 63 (37 C.F.R. § 1.63)
DECLARATION AND POWER OF ATTORNEY
FOR PATENT APPLICATION
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

As below named inventor(s), I/we hereby declare that

This declaration is of the following type:

- | | | |
|---|---------------------------------------|---|
| <input type="checkbox"/> original | <input type="checkbox"/> design | <input type="checkbox"/> supplemental |
| <input checked="" type="checkbox"/> national stage of PCT | | |
| <input type="checkbox"/> divisional | <input type="checkbox"/> continuation | <input type="checkbox"/> continuation-in-part |

My/our residence, post office address and citizenship are as stated below next to my/our name.

I/we believe I/we am/are the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

PROCESS FOR OXIDISING AROMATIC COMPOUNDS ✓

the specification of which (check one)

- ☐ is attached hereto
- ☐ was filed on _____
in the United States Patent and Trademark Office as Application Serial No. _____
and was amended on _____ (if applicable)
- ☐ was described and claimed in PCT International Application No. GB00/02379 ✓
filed on 19 Jun 2000 ✓
and as amended under PCT Article 19 on _____ (if any)

I/we hereby state that I/we have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I/we acknowledge the duty to disclose information which is material to patent ability as defined in 37 C.F.R. § 1.56.

I/we hereby claim foreign priority benefits under 35 U.S.C §119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate or under § 365(a) of any PCT International Application(s) which designated at least one country other than the United States of America, listed below and have also identified below any foreign application for patent or inventor's certificate or PCT International Application having a filing date before that of the application on which priority is claimed:

COMBINED DECLARATION AND POWER OF ATTORNEY**PRIOR FOREIGN/PCT APPLICATION(S) AND ANY PRIORITY CLAIMS UNDER 35 U.S.C. §119**

Application No.	Country	Filing Date	Priority Claimed	
			Yes	No
9914373.7 ✓	GB ✓	18 Jun 1999 ✓	X	
GB00/02379 ✓		19 Jun 2000 ✓		

I/we hereby claim the benefit under 35 U.S.C. § 119(e) of any United States Provisional Application(s) listed below:

UNITED STATES PROVISIONAL APPLICATION(S)

Application No.	Filing Date

I/we hereby claim the benefit under 35 U.S.C. § 120 of any United States Application(s) or § 365(c) of any PCT International Application(s) designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International Application in the manner provided by the first paragraph of 35 U.S.C. § 112, I/we acknowledge the duty to disclose information which is material to patent ability as defined in 37 C.F.R. § 1.56 which became available between the filing date of the prior application and the national PCT international filing date of this application.

PRIOR UNITED STATES/PCT INTERNATIONAL APPLICATION(S)

Application No.	Filing Date	Status (patented, pending/abandoned)
GB00/02379 ✓	19 Jun 2000 ✓	

I/we hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and to transact all business in the U.S. Patent and Trademark Office connected therewith:

Fulbright & Jaworski LLP, 1301 McKinney Suite 5100 Houston Texas 77010-3095 United States of America

I/we hereby declare that all statements made herein of my/our own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C § 1001 and that such willful false statements may jeopardise the validity of the application or any patent issued thereon.

COMBINED DECLARATION AND POWER OF ATTORNEY

Luet wong
Inventors Signature

7/12/2001
Date

1-00 Full name of first/sole inventor Luet Lok WONG
Citizenship the United Kingdom ✓
Residence(City) the United Kingdom
Post Office Address University of Oxford Department of Chemistry
South Parks Road Oxford OX1 3QR the United Kingdom GBX

J. Jones
Inventors Signature

2/1/2002
Date

2-00 Full name of 2nd inventor Jonathan Peter JONES
Citizenship the United Kingdom
Residence(City) the United Kingdom
Post Office Address University of Oxford Department of Chemistry
South Parks Road Oxford OX1 3QR the United Kingdom GBX

I hereby certify that this correspondence is being deposited with the U S Postal Service as Express Mail, Airbill No EU098493435US, in an envelope addressed to Box Sequence, Commissioner for Patents, Washington, DC 20231, on the date shown below

Dated, April 4, 2002

Signature

Linda A Bourg
(Linda A Bourg)

1000 313 730 01101012
13 Rec'd PCT/PTO 01 APR 2002
10/018730

Docket No.: HO-P02353US1
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
Luet L. Wong, et al.

Application No.: 10/018,730
Filed: December 18, 2001

Group Art Unit: N/A

Intl Appln. No.: PCT/GB00/02379
Filed: June 19, 2000

Examiner: Not Yet Assigned

For: PROCESS FOR OXIDISING AROMATIC
COMPOUNDS.

**RESPONSE TO NOTICE TO COMPLY WITH REQUIREMENTS
FOR PATENT APPLICATION CONTAINING NUCLEOTIDE
SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES**

Box Sequence

Commissioner for Patents
Washington, DC 20231

Dear Sir:

This is in response to the Notice To Comply With Requirements For Patent Applications Containing Nucleotide Sequence and/or Amino Acid Sequence Disclosures, dated February 4, 2002. Applicants are submitting their response herewith within the two-month response period set to expire on April 4, 2002.

The U.S. Patent and Trademark Office has taken the position that the previous "Sequence Listing" in computer readable form did not comply with the requirements of Section 1.824 for the stated reason that it contained several raw sequence listing errors. Submitted herewith in connection with the above-referenced patent application is a substitute, computer readable copy of the Sequence Listing that corrects the defects noted in the raw sequence listing. No new matter has been added to the substitute Sequence Listing.

Application No.: 10/018,730


Docket No.: HO-P02353US1

Applicants respectfully request the above-referenced patent application be placed upon the files for examination.

If any fees are due in this matter, please charge our Deposit Account, No. 06-2375, under Order No. 10112404 from which the undersigned is authorized to draw.

Dated: April 4, 2002

Respectfully submitted,

By 
Elizabeth A. Hart, Registered Patent Agent
Registration No.: 50,931
FULBRIGHT & JAWORSKI L.L.P.
1301 McKinney, Suite 5100
Houston, Texas 77010-3095
(713) 651-5698
(713) 651-5246 (Fax)

SEQUENCE LISTING

<110> Luet, Wong
Jonathan, Jones

<120> PROCESS FOR OXIDISING AROMATIC COMPOUNDS

<130> P02353US1 / 10112404 / N.76277B

<140> US 10/018,730

<141> 2001-12-18

<150> GB 9914373.7

<151> 1999-06-18

<150> PCT/GB00/02379

<151> 2000-06-19

<160> 18

<170> PatentIn version 3.1

<210> 1

<211> 1242

<212> DNA

<213> Pseudomonas putida

<400> 1

```
acgactgaaa ccatacaaag caacgccaat cttgcccctc tgccacccca tgtgccagag      60
cacctgggat tgcacttcga catgtacaat ccgtcgaaac tgtctgccgg cgtgcaggag      120
gcctgggacg ttctgcaaga atcaaacgta ccggatctgg tgtggactcg ctgcaacggc      180
ggacactgga tgcacctcg cggccaactg atccgtgagg cctatgaaga ttaccgccac      240
ttttccagcg agtgcccgtt catcctctgt gaagccggcg aagcctacga cttcattccc      300
acctcgatgg atccgcccga gcagcgccag ttctgtgcgc tggccaacca agtggttggc      360
atgccggttg tggataagct ggagaaccgg atccaggagc tggcctgctc gctgatcgag      420
agcctgcgcc cgcaaggaca gtgcaacttc accgaggact acgccgaacc cttcccgata      480
cgcatcttca tgctgctcgc aggtctaccg gaagaagata tcccgcaactt gaaataccta      540
acggatcaga tgacccgtcc ggatggcagc atgaccttcg cagaggccaa ggaggcgctc      600
tacgactatc tgataccgat catcgagcaa cgcaggcaga agccgggaac cgacgctatc      660
agcatcggtt ccaacggcca ggtcaatggg cgaccgatca ccagtgacga agccaagagg      720
atgtgtggcc tggtactggc cggcggcctg gatacgggtg tcaatttcct cagcttcagc      780
atggagttcc tggccaaaag cccggagcat cgccaggagc tgatcgagcg tcccagcgct      840
attccagccg cttgcgagga actactccgg cgcttctcgc tggttgccga tggccgcata      900
ctcacctccg attacgagtt tcatggcgtg caactgaaga aaggtgacca gatcctgcta      960
ccgcagatgc tgtctggcct ggatgagcgc gaaaacgcct gcccgatgca cgtcgacttc     1020
```

ALL INFORMATION CONTAINED HEREIN IS UNCLASSIFIED

```
<210> 2
<211> 414
<212> PRT
<213> Pseudomonas putida
```

<400> 2

Thr Thr Glu Thr Ile Gln Ser Asn Ala Asn Leu Ala Pro Leu Pro Pro
1 5 10 15

His Val Pro Glu His Leu Val Phe Asp Phe Asp Met Tyr Asn Pro Ser
20 25 30

Asn Leu Ser Ala Gly Val Gln Glu Ala Trp Ala Val Leu Gln Glu Ser
35 40 45

Asn Val Pro Asp Leu Val Trp Thr Arg Cys Asn Gly Gly His Trp Ile
50 55 60

Ala Thr Arg Gly Gln Leu Ile Arg Glu Ala Tyr Glu Asp Tyr Arg His
65 70 75 80

Phe Ser Ser Glu Cys Pro Phe Ile Pro Arg Glu Ala Gly Glu Ala Tyr
85 90 95

Asp Phe Ile Pro Thr Ser Met Asp Pro Pro Glu Gln Arg Gln Phe Arg
100 105 110

Ala Leu Ala Asn Gln Val Val Gly Met Pro Val Val Asp Lys Leu Glu
115 120 125

Asn Arg Ile Gln Glu Leu Ala Cys Ser Leu Ile Glu Ser Leu Arg Pro
130 135 140

Gln Gly Gln Cys Asn Phe Thr Glu Asp Tyr Ala Glu Pro Phe Pro Ile
145 150 155 160

Arg Ile Phe Met Leu Leu Ala Gly Leu Pro Glu Glu Asp Ile Pro His
165 170 175

<211> 3150
 <212> DNA
 <213> Bacillus megaterium

<400> 3
 atgacaatta aagaaatgcc tcagccaaaa acgtttggag agcttaaaaa tttaccgtta 60
 ttaaacacag ataaaccggt tcaagctttg atgaaaattg cggatgaatt aggagaaatc 120
 tttaaattcg aggcgcctgg tcgtgtaacg cgctacttat caagtcagcg tctaattaaa 180
 gaagcatgcg atgaatcacg ctttgataaa aacttaagtc aagcgcttaa atttgtaacgt 240
 gattttgcag gagacgggtt atttacaagc tggacgcagc aaaaaaattg gaaaaaagcg 300
 cataatatct tacttccaag cttcagtcag caggcaatga aaggctatca tgcgatgatg 360
 gtcgatatcg ccgtgcagct tgttcaaaag tgggagcgtc taaatgcaga tgagcatatt 420
 gaagtaccgg aagacatgac acgtttaacg cttgatacaa ttggtctttg cggctttaac 480
 tatcgcttta acagctttta ccgagatcag cctcatccat ttattacaag tatgggtccgt 540
 gcactggatg aagcaatgaa caagctgcag cgagcaaatc cagacgaccc agcttatgat 600
 gaaaacaagc gccagtttca agaagatatc aagggtgatga acgacctagt agataaaatt 660
 attgcagatc gcaaagcaag cgggtgaacaa agcgatgatt tattaacgca tatgctaaac 720
 ggaaaagatc cagaaacggg tgagccgctt gatgacgaga acattcgcta tcaaattatt 780
 acattcttaa ttgcgggaca cgaaacaaca agtgggtcttt tatcatttgc gctgtatttc 840
 ttagtgaaaa atccacatgt attacaaaaa gcagcagaag aagcagcacg agttctagta 900
 gatcctgctc caagctacaa acaagtcaaa cagcttaaat atgtcggcat ggtcttaaac 960
 gaagcgctgc gcttatggcc aactgctcct gcgttttccc tatatgcaaa agaagatacg 1020
 gtgcttggag gagaatatcc tttagaaaaa ggcgacgaac taatggttct gattcctcag 1080
 cttcacccgtg ataaaaacaat ttggggagac gatgtggaag agttccgtcc agagcggttt 1140
 gaaaatccaa gtgcgattcc gcagcatgcg tttaaaccgt ttggaaacgg tcagcgtgcg 1200
 tgtatcggtc agcagttcgc tcttcatgaa gcaacgctgg tacttggtat gatgctaaaa 1260
 cactttgact ttgaagatca taaaactac gagctggata ttaaagaaac tttaacgtta 1320
 aaacctgaag gctttgtggt aaaagcaaaa tcgaaaaaaa ttccgcttgg cggtatctct 1380
 tcacctagca ctgaacagtc tgccaaaaaa gcacgcaaaa aggcagaaaa cgctcataat 1440
 acgccgctgc ttgtgctata cggttcaa atgggaacag ctgaaggaac ggcgcgtgat 1500
 ttagcagata ttgcaatgag caaaggattt gcaccgcagg tcgcaacgct tgattcacac 1560
 gccggaaatc ttccgcgcga aggagctgta ttaattgtaa cggcgtctta taacgggtcat 1620
 ccgcctgata acgcaaagca atttgctgac tggtagacc aagcgtctgc tgatgaagta 1680

aaaggcggttc gctactccgt atttgatgc ggcgataaaa actgggctac tacgtatcaa 1740
aaagtgcctg cttttatcga tgaaacgctt gccgctaaag gggcagaaaa catcgctgac 1800
cgcggtgaag cagatgcaag cgacgacttt gaaggcacat atgaagaatg gcgtgaacat 1860
atgtggagtg acgtagcagc ctactttaac ctgcacattg aaaacagtga agataataaa 1920
tctactcttt cacttcaatt tgtcgacagc gccgcgata tgccgcttgc gaaaatgcac 1980
ggtgcgtttt caacgaacgt cgtagcaagc aaagaacttc aacagccagg cagtgcacga 2040
agcacgcgac atcttgaaat tgaacttcca aaagaagctt cttatcaaga aggagatcat 2100
ttaggtgtta ttctcgcga ctatgaagga atagtaaacc gtgtaacagc aagggttcggc 2160
ctagatgcat cacagcaaat ccgtctggaa gcagaagaag aaaaattagc tcatttgcca 2220
ctcgctaaaa cagtatccgt agaagagctt ctgcaatacg tggagcttca agatcctgtt 2280
acgcgcacgc agcttcgcgc aatggctgct aaaacggtct gcccgccgca taaagtagag 2340
cttgaagcct tgcttgaaaa gcaagcctac aaagaacaag tgctggcaaa acgtttaaca 2400
atgcttgaac tgcttgaaaa ataccggcg tgtgaaatga aattcagcga atttatcgcc 2460
cttctgccaa gcatacgccc gcgctattac tcgatttctt catcacctcg tgtcgatgaa 2520
aaacaagcaa gcatcacggt cagcgttgtc tcaggagaag cgtggagcgg atatggagaa 2580
tataaaggaa ttgcgtcgaa ctatcttgcc gagctgcaag aaggagatac gattacgtgc 2640
tttatttcca caccgcagtc agaatttacg ctgccaaaag acctgaaac gccgcttate 2700
atggctcgac cggaacagg cgtcgcgccg ttagaggct ttgtgcaggc gcgcaaacag 2760
ctaaaagaac aaggacagtc acttgagaa gcacatttat acttcggctg ccgttcacct 2820
catgaagact atctgtatca agaagagctt gaaaacgccc aaagcgaagg catcattacg 2880
cttcataccg ctttttctcg catgccaaat cagccgaaaa catacgttca gcacgtaatg 2940
gaacaagacg gcaagaaatt gattgaactt cttgatcaag gagcgcaactt ctatatttgc 3000
ggagacggaa gccaaatggc acctgccgtt gaagcaacgc ttatgaaaag ctatgctgac 3060
gttcaccaag tgagtgaagc agacgctcgc ttatggctgc agcagctaga agaaaaaggc 3120
cgatacgcaa aagacgtgtg ggctgggtaa 3150

<210> 4
<211> 1049
<212> PRT
<213> Bacillus megaterium

<400> 4

Met Thr Ile Lys Glu Met Pro Gln Pro Lys Thr Phe Gly Glu Leu Lys
1 5 10 15

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) = \frac{\partial L}{\partial x}, \quad \text{(a)} \quad \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{y}} \right) = \frac{\partial L}{\partial y}$$

Asn Leu Pro Leu Leu Asn Thr Asp Lys Pro Val Gln Ala Leu Met Lys
20 25 30

Ile Ala Asp Glu Leu Gly Glu Ile Phe Lys Phe Glu Ala Pro Gly Arg
35 40 45

Val Thr Arg Tyr Leu Ser Ser Gln Arg Leu Ile Lys Glu Ala Cys Asp
50 55 60

Glu Ser Arg Phe Asp Lys Asn Leu Ser Gln Ala Leu Lys Phe Val Arg
65 70 75 80

Asp Phe Ala Gly Asp Gly Leu Phe Thr Ser Trp Thr His Glu Lys Asn
85 90 95

Trp Lys Lys Ala His Asn Ile Leu Leu Pro Ser Phe Ser Gln Gln Ala
100 105 110

Met Lys Gly Tyr His Ala Met Met Val Asp Ile Ala Val Gln Leu Val
115 120 125

Gln Lys Trp Glu Arg Leu Asn Ala Asp Glu His Ile Glu Val Pro Glu
130 135 140

Asp Met Thr Arg Leu Thr Leu Asp Thr Ile Gly Leu Cys Gly Phe Asn
145 150 155 160

Tyr Arg Phe Asn Ser Phe Tyr Arg Asp Gln Pro His Pro Phe Ile Thr
165 170 175

Ser Met Val Arg Ala Leu Asp Glu Ala Met Asn Lys Leu Gln Arg Ala
180 185 190

Asn Pro Asp Asp Pro Ala Tyr Asp Glu Asn Lys Arg Gln Phe Gln Glu
195 200 205

Asp Ile Lys Val Met Asn Asp Leu Val Asp Lys Ile Ile Ala Asp Arg
210 215 220

Lys Ala Ser Gly Glu Gln Ser Asp Asp Leu Leu Thr His Met Leu Asn
225 230 235 240

Gly Lys Asp Pro Glu Thr Gly Glu Pro Leu Asp Asp Glu Asn Ile Arg
245 250 255

[illegible]

Leu Leu Ser Phe Ala Leu Tyr Phe Leu Val Lys Asn Pro His Val Leu
275 280 285

Gln Lys Ala Ala Glu Glu Ala Ala Arg Val Leu Val Asp Pro Ala Pro
290 295 300

Ser Tyr Lys Gln Val Lys Gln Leu Lys Tyr Val Gly Met Val Leu Asn
305 310 315 320

Glu Ala Leu Arg Leu Trp Pro Thr Ala Pro Ala Phe Ser Leu Tyr Ala
325 330 335

Lys Glu Asp Thr Val Leu Gly Gly Glu Tyr Pro Leu Glu Lys Gly Asp
340 345 350

Glu Leu Met Val Leu Ile Pro Gln Leu His Arg Asp Lys Thr Ile Trp
355 360 365

Gly Asp Asp Val Glu Glu Phe Arg Pro Glu Arg Phe Glu Asn Pro Ser
370 375 380

Ala Ile Pro Gln His Ala Phe Lys Pro Phe Gly Asn Gly Gln Arg Ala
385 390 395 400

Cys Ile Gly Gln Gln Phe Ala Leu His Glu Ala Thr Leu Val Leu Gly
405 410 415

Met Met Leu Lys His Phe Asp Phe Glu Asp His Thr Asn Tyr Glu Leu
420 425 430

Asp Ile Lys Glu Thr Leu Thr Leu Lys Pro Glu Gly Phe Val Val Lys
435 440 445

Ala Lys Ser Lys Lys Ile Pro Leu Gly Gly Ile Pro Ser Pro Ser Thr
450 455 460

Glu Gln Ser Ala Lys Lys Ala Arg Lys Lys Ala Glu Asn Ala His Asn
465 470 475 480

Thr Pro Leu Leu Val Leu Tyr Gly Ser Asn Met Gly Thr Ala Glu Gly
485 490 495

Thr Ala Arg Asp Leu Ala Asp Ile Ala Met Ser Lys Gly Phe Ala Pro

	500		505		510
Gln Val Ala Thr Leu Asp Ser His Ala Gly Asn Leu Pro Arg Glu Gly	515		520		525
Ala Val Leu Ile Val Thr Ala Ser Tyr Asn Gly His Pro Pro Asp Asn	530		535		540
Ala Lys Gln Phe Val Asp Trp Leu Asp Gln Ala Ser Ala Asp Glu Val	545		550		555
Lys Gly Val Arg Tyr Ser Val Phe Gly Cys Gly Asp Lys Asn Trp Ala	565		570		575
Thr Thr Tyr Gln Lys Val Pro Ala Phe Ile Asp Glu Thr Leu Ala Ala	580		585		590
Lys Gly Ala Glu Asn Ile Ala Asp Arg Gly Glu Ala Asp Ala Ser Asp	595		600		605
Asp Phe Glu Gly Thr Tyr Glu Glu Trp Arg Glu His Met Trp Ser Asp	610		615		620
Val Ala Ala Tyr Phe Asn Leu Asp Ile Glu Asn Ser Glu Asp Asn Lys	625		630		635
Ser Thr Leu Ser Leu Gln Phe Val Asp Ser Ala Ala Asp Met Pro Leu	645		650		655
Ala Lys Met His Gly Ala Phe Ser Thr Asn Val Val Ala Ser Lys Glu	660		665		670
Leu Gln Gln Pro Gly Ser Ala Arg Ser Thr Arg His Leu Glu Ile Glu	675		680		685
Leu Pro Lys Glu Ala Ser Tyr Gln Glu Gly Asp His Leu Gly Val Ile	690		695		700
Pro Arg Asn Tyr Glu Gly Ile Val Asn Arg Val Thr Ala Arg Phe Gly	705		710		715
Leu Asp Ala Ser Gln Gln Ile Arg Leu Glu Ala Glu Glu Glu Lys Leu	725		730		735
Ala His Leu Pro Leu Ala Lys Thr Val Ser Val Glu Glu Leu Leu Gln	740		745		750

Tyr Val Glu Leu Gln Asp Pro Val Thr Arg Thr Gln Leu Arg Ala Met
755 760 765

Ala Ala Lys Thr Val Cys Pro Pro His Lys Val Glu Leu Glu Ala Leu
770 775 780

Leu Glu Lys Gln Ala Tyr Lys Glu Gln Val Leu Ala Lys Arg Leu Thr
785 790 795 800

Met Leu Glu Leu Leu Glu Lys Tyr Pro Ala Cys Glu Met Lys Phe Ser
805 810 815

Glu Phe Ile Ala Leu Leu Pro Ser Ile Arg Pro Arg Tyr Tyr Ser Ile
820 825 830

Ser Ser Ser Pro Arg Val Asp Glu Lys Gln Ala Ser Ile Thr Val Ser
835 840 845

Val Val Ser Gly Glu Ala Trp Ser Gly Tyr Gly Glu Tyr Lys Gly Ile
850 855 860

Ala Ser Asn Tyr Leu Ala Glu Leu Gln Glu Gly Asp Thr Ile Thr Cys
865 870 875 880

Phe Ile Ser Thr Pro Gln Ser Glu Phe Thr Leu Pro Lys Asp Pro Glu
885 890 895

Thr Pro Leu Ile Met Val Gly Pro Gly Thr Gly Val Ala Pro Phe Arg
900 905 910

Gly Phe Val Gln Ala Arg Lys Gln Leu Lys Glu Gln Gly Gln Ser Leu
915 920 925

Gly Glu Ala His Leu Tyr Phe Gly Cys Arg Ser Pro His Glu Asp Tyr
930 935 940

Leu Tyr Gln Glu Glu Leu Glu Asn Ala Gln Ser Glu Gly Ile Ile Thr
945 950 955 960

Leu His Thr Ala Phe Ser Arg Met Pro Asn Gln Pro Lys Thr Tyr Val
965 970 975

Gln His Val Met Glu Gln Asp Gly Lys Lys Leu Ile Glu Leu Leu Asp
980 985 990

Gln Gly Ala His Phe Tyr Ile Cys Gly Asp Gly Ser Gln Met Ala Pro
 995 1000 1005

Ala Val Glu Ala Thr Leu Met Lys Ser Tyr Ala Asp Val His Gln
 1010 1015 1020

Val Ser Glu Ala Asp Ala Arg Leu Trp Leu Gln Gln Leu Glu Glu
 1025 1030 1035

Lys Gly Arg Tyr Ala Lys Asp Val Trp Ala Gly
 1040 1045

<210> 5
 <211> 51
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer

<220>
 <221> misc_feature
 <222> (1)..(51)
 <223>

<400> 5
 gagattaaga attcataaac acatgggagt gcgtgccata tgaacgcaaa c 51

<210> 6
 <211> 7
 <212> PRT
 <213> Artificial sequence

<220>
 <223> Linker

<220>
 <221> MISC_FEATURE
 <222> (1)..(7)
 <223>

<400> 6
 Thr Asp Gly Gly Ser Ser Ser
 1 5

<210> 7
 <211> 36
 <212> DNA
 <213> Artificial sequence

<220>

100193730 040102

 $\langle 220 \rangle$

<221> misc feature

<222> (1) . . (36)

<223>

<400> 7

gaactgagta gtgccactga cggaggatcc tcatcg

36

<210> 8

$\langle 211 \rangle$	5
-----------------------	---

<212> PRT

<213> Artificial sequence

 $\langle 220 \rangle$

<223> Desired coding sequence

 $\langle 220 \rangle$

<221> MISC FEATURE

$\langle 222 \rangle$ (1) . . (5)

<223>

<400> 8

Thr Asp Gly Gly Ser

1 5

<210> 9

<211> 36

<212> DNA

<213> Artificial sequence

 $\langle 220 \rangle$

<223> Primer

 $\langle 220 \rangle$

<221> misc feature

$$\langle 222 \rangle \quad (1) \cdot \cdot (36)$$

<223>

<400> 9

cgatgaggat cctccgtcag tggcactact cagttc

36

 $\langle 210 \rangle$ 10

<211> 36

<212> DNA

<213> Artificial sequence

 $\langle 220 \rangle$

<223> Desired coding sequence

 $\langle 220 \rangle$

```
<221> misc_feature
```

 $\langle 222 \rangle \quad (1) \cdot \bar{\cdot} (36)$

<223>

<400> 10
tcatcgggat cctcatcgat gtctaaagta gtgtat

36

<210> 11
<211> 4
<212> PRT
<213> Artificial sequence

<220>
<223> Desired coding sequence

<220>
<221> MISC_FEATURE
<222> (1)..(4)
<223>

<400> 11
Gly Ser Ser Ser
1

<210> 12
<211> 52
<212> DNA
<213> Artificial sequence

<220>
<223> Desired coding sequence

<220>
<221> misc_feature
<222> (1)..(52)
<223>

<400> 12
cccgataggc aatggtaatc atcgggagtc tagagcatcg aagctttcat cg

52

<210> 13
<211> 52
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<220>
<221> misc_feature
<222> (1)..(52)
<223>

<400> 13
cgatgaaagc ttcgatgctc tagactcccg atgattacca ttgcctatcg gg

52

